

*Maria Esther de Noronha Fonseca
Leonardo Silva Boiteux*

1. Introdução

O melhoramento genético (em geral) e o melhoramento genético para resistência a patógenos e pragas (em particular) têm viabilizado, por décadas, o cultivo de plantas de forma mais sustentável (Zamir, 2001). Entre as ferramentas disponíveis para os programas de melhoramento genético para resistência a patógenos estão os métodos de melhoramento envolvendo cruzamentos controlados e bioensaios com inoculação controlada para a seleção fenotípica de plantas individuais que combinem resistência contra uma doença alvo e características agrônômicas de interesse para o consumidor final (Boiteux et al., 2016). Outro aspecto de infraestrutura importante é a disponibilidade de bancos de germoplasma contendo a mais ampla diversidade genética e extensas coleções de mutantes, incluindo fatores de resistência aos principais patógenos (Emmanuel & Levy, 2002). Os cruzamentos controlados para incorporação de fatores de resistência envolvem, em geral, cultivares comerciais (ou linhagens elite) e acessos de espécies silvestres (= **fontes de resistência**) de onde os genes de interesse vão ser **introgrididos**. Desta forma, é necessário, após os cruzamentos interespecíficos iniciais, tecnologias capazes de “recuperar” de maneira eficiente o genoma da linhagem elite, contendo agora o novo fator de resistência. Além disso, o desenvolvimento de cultivares com resistência a doenças via melhoramento genético é, necessariamente, um processo contínuo uma vez que novos patógenos e novas variantes de patógenos surgem (ou podem ser introduzidos) a todo momento em uma dada região geográfica. Esse aspecto dinâmico do processo de incorporação e piramidização de novos fatores requer rapidez e precisão, especialmente no caso de hortaliças que são afetadas por uma assombrosa gama de patógenos em constante evolução. Neste cenário, a utilização de **seleção assistida por marcadores moleculares (SAM)** tem se tornado um componente metodológico crucial, apresentando uma perfeita harmonia com todos os métodos clássicos de melhoramento genético, permitindo uma seleção mais robusta (Ferreira & Grattapaglia, 1998). Por sua vez, os marcadores moleculares e o desenvolvimento de mapas genéticos ultradensos conduzem para a localização física e a posterior descoberta de novos genes de resistência bem como seus genes reguladores. A localização cromossômica e a descoberta de novos fatores de resistência têm se constituído em atividades essenciais na aplicação de novas estratégias de melhoramento via biotecnologia tais como transformação genética e edição genômica (Rothan et al., 2019). Todas estas tecnologias avançadas são mais eficientes com o conhecimento preciso da região do genoma a ser editada ou a ser mobilizada. Essa informação é proveniente, em última análise, dos estudos genéticos conduzidos com o auxílio dos marcadores moleculares. Neste capítulo será apresentan-

do um breve histórico das tecnologias que permitiram o desenvolvimento dos principais sistemas de marcadores moleculares bem como um conjunto de exemplos da aplicação de SAM no melhoramento genético para resistência a patógenos usando o tomateiro (*Solanum lycopersicum* L.) como planta modelo.

2. Evoluções Tecnológicas da Era Genômica que Permitiram a Geração dos Modernos Sistemas de Marcadores Moleculares

2.1. Elucidação dos mecanismos de hereditariedade e da estrutura do DNA

A elucidação da estrutura do DNA por Watson & Crick (1953), a definição de seu papel crucial na transmissão de caracteres e o estabelecimento definitivo do “dogma central” da genética (**Figura 1**) possibilitaram inúmeros e consistentes avanços no melhoramento vegetal. Esse conjunto de conhecimentos genéticos claramente indicou que dissecando a estrutura do DNA podemos decifrar os mecanismos de transmissão de todos os caracteres (qualitativos e quantitativos) de interesse, incluindo fatores de resistência ou suscetibilidade a doenças em plantas.

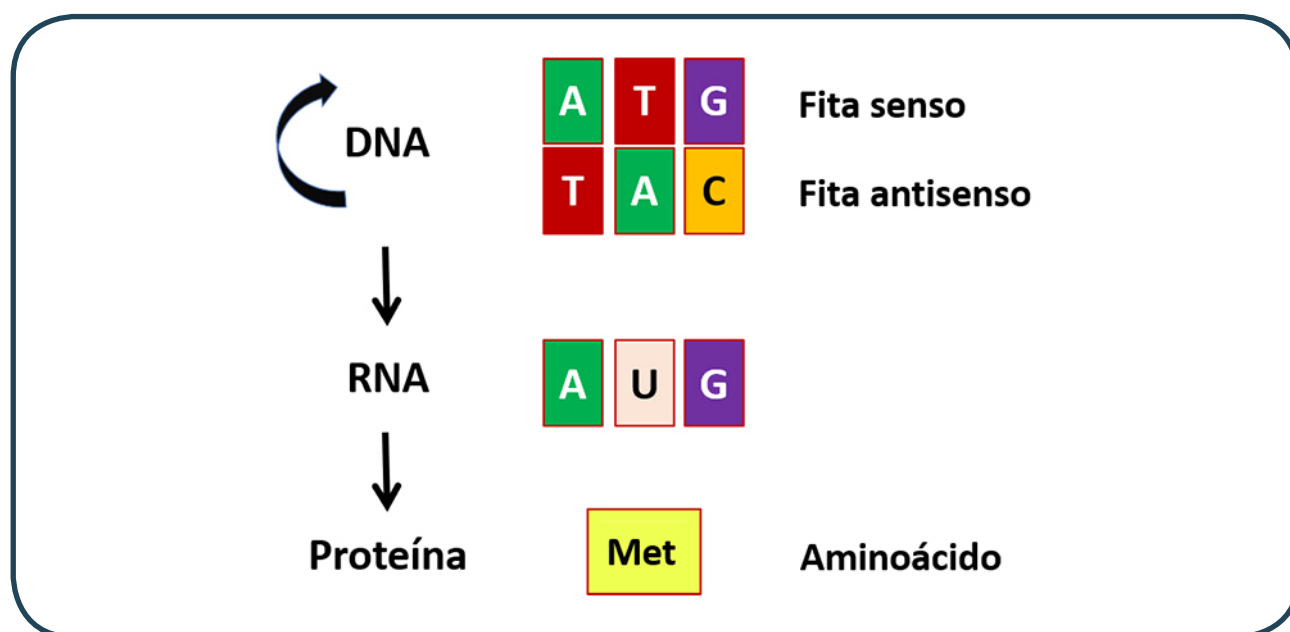


Figura 1. Diagrama do fluxo de informação genética (dogma central da genética). A molécula de DNA carrega toda a informação para própria replicação e para a sua transcrição em RNA. O RNA, de acordo com o código genético, é traduzido em aminoácidos que formam as proteínas. Na ilustração, o códon ATG na fita senso do DNA é transcrito em AUG e finalmente traduzido no aminoácido metionina.

2.2. Metodologias de sequenciamento da molécula do DNA

A definição da sequência de bases do DNA e a identificação das diferenças genéticas que geram variabilidade alélica bem como expressão de distintos caracteres formam as bases do moderno melhoramento genético vegetal. A elucidação das funções do DNA estimulou a busca por técnicas de sequenciamento mais precisas, robustas e baratas para a comunidade científica. O método desenvolvido por Sanger et al. (1977) foi o primeiro grande passo que permitiu o sequenciamento em mais ampla escala (**Figura 2**). Esse método baseia-se no estabelecimento de reações bioquímicas de síntese de DNA contendo uma combinação de deoxiribonucleotídeos normais (dATP, dGTP, dCTP e dTTT) e alguns dideoxiribonucleotídeos com bases modificadas (ddATP, ddGTP, ddCTP e ddTTT) que provocam uma interrupção na amplificação do DNA. Essas bases modificadas são marcadas permitindo sua posterior detecção. Esse procedimento resulta em uma amplitude de fragmentos de vários tamanhos que podem ser separados por eletroforese e detectados em sequenciadores automáticos (**Figura 2**). Atualmente o método de Sanger é considerado como “sequenciamento de primeira geração” e se constitui no padrão de validação de mutações observadas nas tecnologias de nova geração. O método Sanger, no entanto, demonstrou ter limitações de escala, evidenciadas em projetos mais ambiciosos de sequenciamento de genomas completos e de várias amostras (ou variantes) de um mesmo organismo em paralelo (ao mesmo tempo). Essa ampliação de escala foi obtida através de novas tecnologias coletivamente denominadas como **NGS (= Next Generation Sequencing)**. NGS é um termo geral para denominar os métodos de sequenciamento em larga escala ou “em paralelo”. Em geral, o NGS é realizado em moldes imobilizados em membranas e a adição de bases é monitorada. A técnica de obtenção das sequências, no entanto, varia entre os vários modelos de sequenciadores de nova geração (Linnarsson, 2010). Um esquema simplificado da estratégia de NGS é ilustrado na **Figura 3**. As principais limitações das tecnologias NGS são o custo por fragmento e os tamanhos dos fragmentos obtidos. As principais plataformas (Illumina, Roche, SoLiD) disponibilizam fragmentos curtos (entre 30 e 400 pb) se comparados aos obtidos por Sanger (entre 500 e 1000 pb). Novas plataformas (PacBio e Nanopore) e novos protocolos (Illumina, Roche, SoLiD) geram fragmentos maiores, mas ainda com custos elevados. A grande limitação dos fragmentos curtos é a dificuldade de montagem dos genomas se realizados “de novo”, ou seja, sem um **genoma de referência** como molde (**Figura 3**). Por esta razão, a disponibilidade de genomas prontos, anotados e montados, como no caso do tomateiro, é de extrema utilidade para o melhoramento genético. Além disso, é importante que sejam obtidos vários fragmentos de DNA que se sobreponham, permitindo a montagem de longos “contigs”, resultando em uma boa “cobertura” do genoma e evitando lacunas de informação de sequência nos diferentes cromossomos ou regiões genômicas. Com o advento das técnicas de NGS os genomas de referência de várias culturas passaram a estar disponíveis, facilitando o chamado **resequenciamento genômico**, importante ferramenta para estudo da diversidade e detecção de diferenças genéticas entre acessos de uma dada espécie. Com estes moldes ou genomas de referência, a montagem dos cromossomos via resequenciamento se torna exequível mesmo com fragmentos pequenos, embora possam persistir alguns problemas em algumas regiões específicas

que podem não apresentar “cobertura” suficiente. Dois exemplos de resequenciamento do tomateiro são os projetos “BGI tomato 360 genomes” (Lin et al., 2014) e “100 tomato genome resequencing project” (Tomato Genome Sequencing Consortium, 2012; 2014). As informações sobre o genoma do tomateiro e de outras espécies cultivadas e selvagens da família Solanácea estão disponíveis no portal da SolGenomics (www.solgenomics.net).

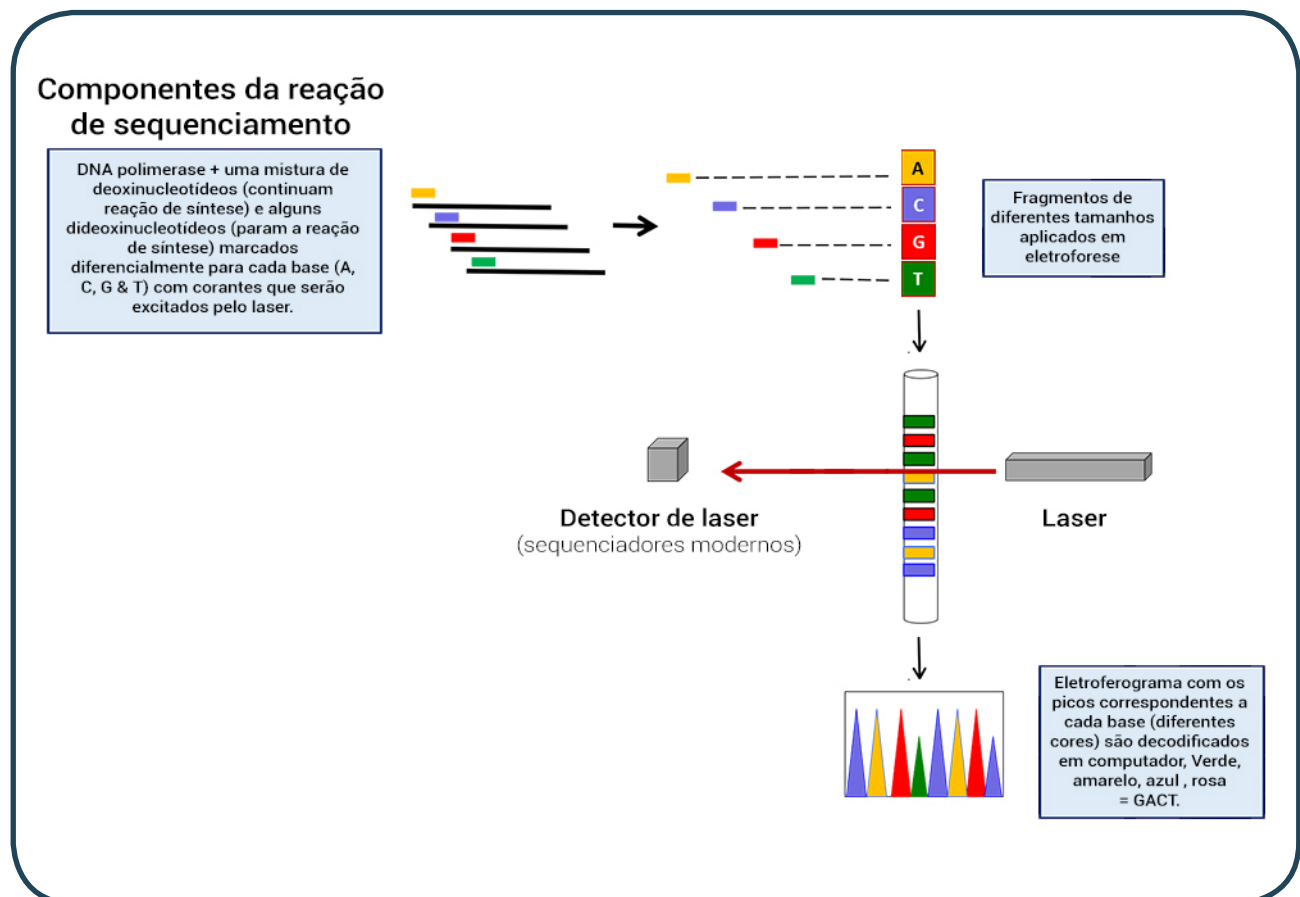


Figura 2. Diagrama ilustrando os principais componentes do método de sequenciamento de Sanger em um sequenciador automático. O conjunto de amplicons de diferentes tamanhos e marcados diferencialmente (no exemplo, amarelo = A, azul = C, vermelho = G e verde = T) são obtidos via PCR, separados em capilares e cada um é sequencialmente detectado e sua informação ordenada por tamanho.

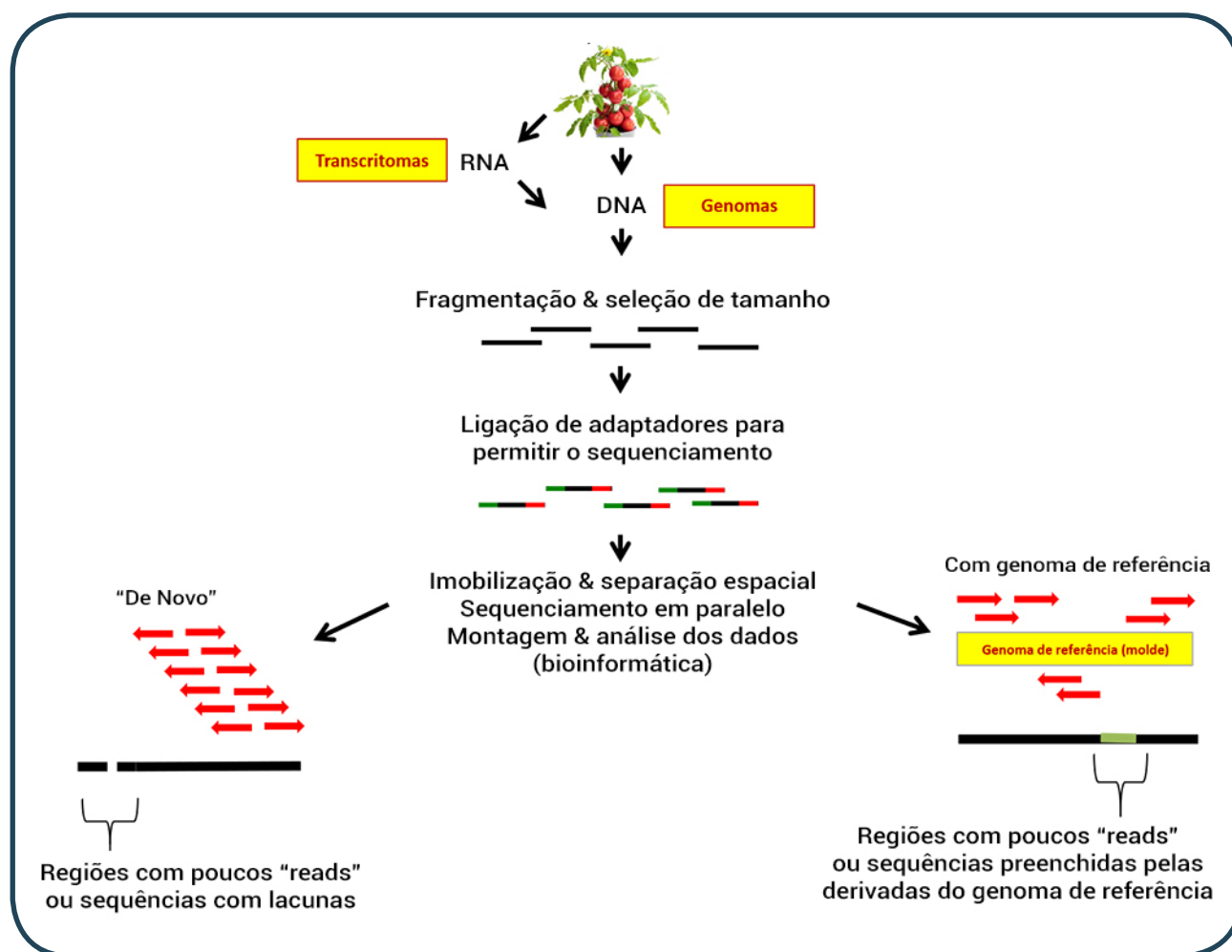


Figura 3. Esquema de NGS (Next Generation Sequencing). NGS é um termo geral para denominar os métodos de sequenciamento em larga escala ou "em paralelo". Em geral, o NGS é realizado em moldes imobilizados em membranas e a adição de bases é monitorada. As principais plataformas de NGS (Illumina, Roche, SoLiD) disponibilizam fragmentos curtos (entre 30 e 400 pares de bases). A grande limitação dos fragmentos curtos é a dificuldade de montagem dos genomas se realizados "de novo", ou seja, sem um genoma de referência como molde. Novas plataformas (Nanopore, PacBio) ou novos protocolos (Illumina, Roche, SoLiD) têm sido desenvolvidos para solucionar essa limitação do NGS.

2.3. Tecnologia do DNA recombinante & engenharia genética

Duas descobertas permitiram o desenvolvimento da chamada **engenharia genética**. Primeiro a confirmação do DNA como a molécula unificadora de todos os sistemas biológicos. Todos os seres vivos transferem informações utilizando DNA, RNA e proteínas, ou seja, o código genético é universal. Outra descoberta metodológica importante foi a das **enzimas de restrição**, que são normalmente encontradas em bactérias e que apresentam a propriedade de clivar o DNA em sequências palindrômicas específicas. Esses avanços permitiram a inserção de fragmentos específicos de DNA em vetores bacterianos e o intercâmbio destes segmentos gênicos entre diferentes espécies e mesmo entre diferentes reinos. O domínio inicial destas tecnologias permitiu o posterior estabelecimento de todas as técnicas básicas de engenharia genética,

tais como: manutenção do gene de interesse em plasmídeos de bactérias (clonagem), inserção de genes de plantas (e de outros organismos) em *Agrobacterium* e posterior transferência desses genes para o genoma de plantas e, mais recentemente, a inserção em vetores para silenciamento gênico mediados por endonucleases (ZFNs e TALENs) ou CRISPR Cas9 (Rothan et al., 2019).

2.4. PCR (Polymerase Chain Reaction)

PCR é técnica que permite a amplificação de segmentos específicos de DNA tendo suas bordas definidas por sequências com as quais primers anelam para dar início ao processo de replicação. Cópias em série das fitas de DNA são produzidas *in vitro* até que seja obtida a quantidade desejada da região amplificada da região alvo do genoma definida pelos primers (**Figura 4**). Utiliza-se para tal uma enzima do tipo polimerase com propriedades únicas de termo estabilidade (= enzima resiste aos vários ciclos de desnaturação de DNA), permitindo a cópia em série do DNA sem reposição de enzima (Mullis & Faloona, 1987). A técnica de PCR permitiu superar a grande limitação que a disponibilidade de pequenas quantidades de DNA impunha para a pesquisa genética. Com o advento do PCR, mesmas purificações exíguas de DNA puderam ser amplificadas e manipuladas (Mullis & Faloona, 1987). Os aparelhos de PCR são programados para alterar as temperaturas da reação rapidamente (em torno de 94 °C para desnaturação ou separação das fitas de DNA, 37-68 °C para anelamento dos primers, 68-72 °C para síntese de novas fitas com o auxílio das polimerases termoestáveis). Desta forma, em poucos ciclos (em geral 30 ciclos) e pouco tempo (entre 1 e 2 horas) o DNA é amplificado exponencialmente e são obtidas quantidades suficientes para se efetuar a maioria das análises de rotina no melhoramento genético vegetal (exemplo: identificação de marcadores, sequenciamento e clonagem). Em muitos protocolos, o processo de amplificação de mais que uma região alvo pode ser obtida desde que as temperaturas de anelamento dos distintos primers sejam idênticas ou muito próximas, permitindo o estabelecimento do chamado sistema de **PCR multiplex**. Uma modalidade mais recente e refinada desta tecnologia é a chamada **PCR em tempo real (quantitative PCR ou qPCR)**. Nesta técnica, uma nova geração de aparelhos permite o monitoramento do progresso da amplificação dos fragmentos de DNA que são continuamente detectados pela presença de marcação fluorescente diferenciada (Rebrikov & Trofimov, 2006), permitindo identificar as diferentes variantes alélicas pela marcação de iniciadores com diferentes fluoróforos. O sistema facilita o monitoramento da real quantidade inicial do molde de DNA genômico ou derivado de RNA (DNA complementar ou cDNA) e também permite distinguir mutações de acordo com a curva de desnaturação dos amplicons (Novais et al., 2004). No aspecto relacionado ao desenvolvimento de marcadores moleculares, o PCR em tempo real permitiu ampliar a geração de novos sistemas de detecção e análise de polimorfismos de base única (= **Single Nucleotide Polymorphisms – SNPs**), o que anteriormente seria possível apenas por sequenciamento ou pela fortuita presença de uma mudança de nucleotídeo capaz de gerar um sítio de clivagem para alguma enzima de restrição.

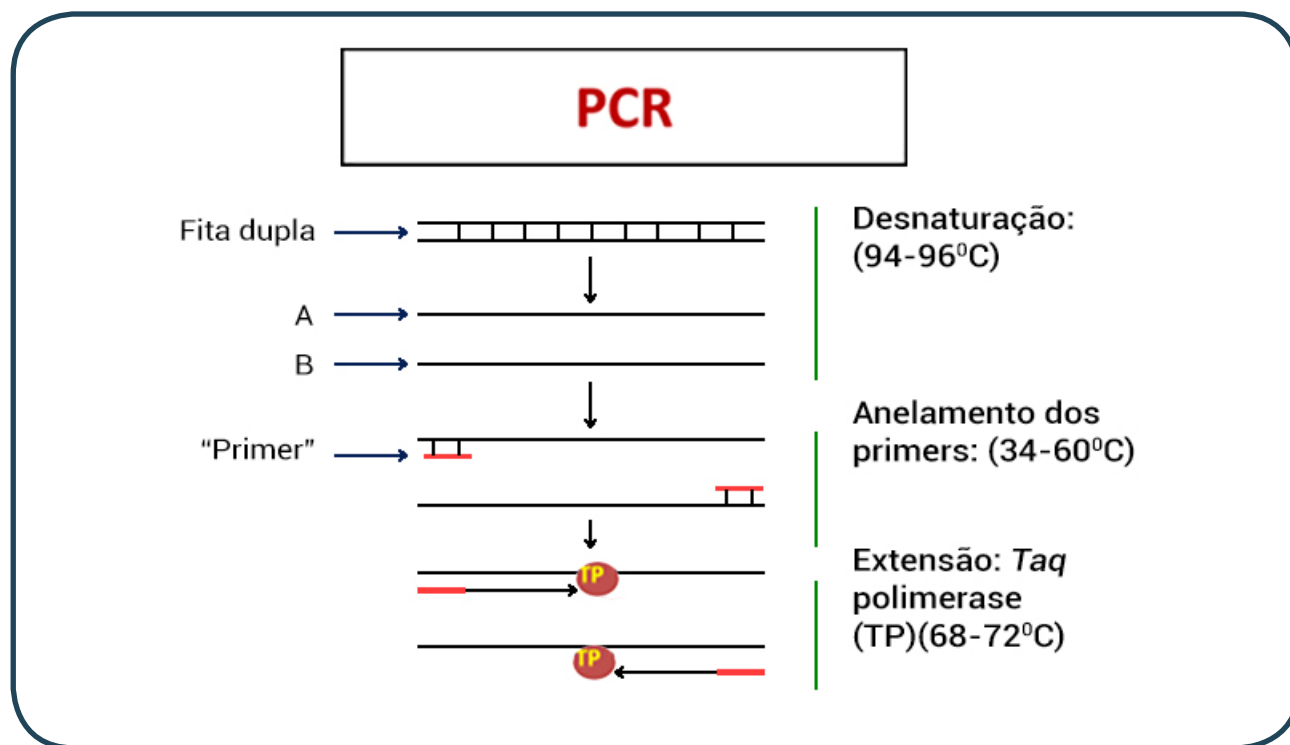


Figura 4. Esquema geral da PCR (Polymerase Chain Reaction). O processo de replicação de DNA é obtido de forma contínua em ciclos de reação que envolvem as etapas de separação ou desnaturaçãõ do DNA fita dupla (94-96 °C), anelamento dos 'primers' ou iniciadores e extensão de novas fitas, realizado com o emprego de enzimas polimerases termoestáveis (exemplo: *Taq* polimerase).

Importância dos marcadores moleculares no melhoramento genético de nova geração

Os marcadores moleculares têm um papel central no moderno melhoramento genético porque interligam as várias áreas de conhecimento fornecendo informações essenciais para a busca de novos genes e posterior utilização em sistemas de transformação genética (transgenia e cisgenia) e edição gênica e, ao mesmo tempo, permitindo que informações disponíveis possam ser diretamente aplicadas em sistemas de SAM e na geração de novas cultivares e híbridos (**Figura 5**).

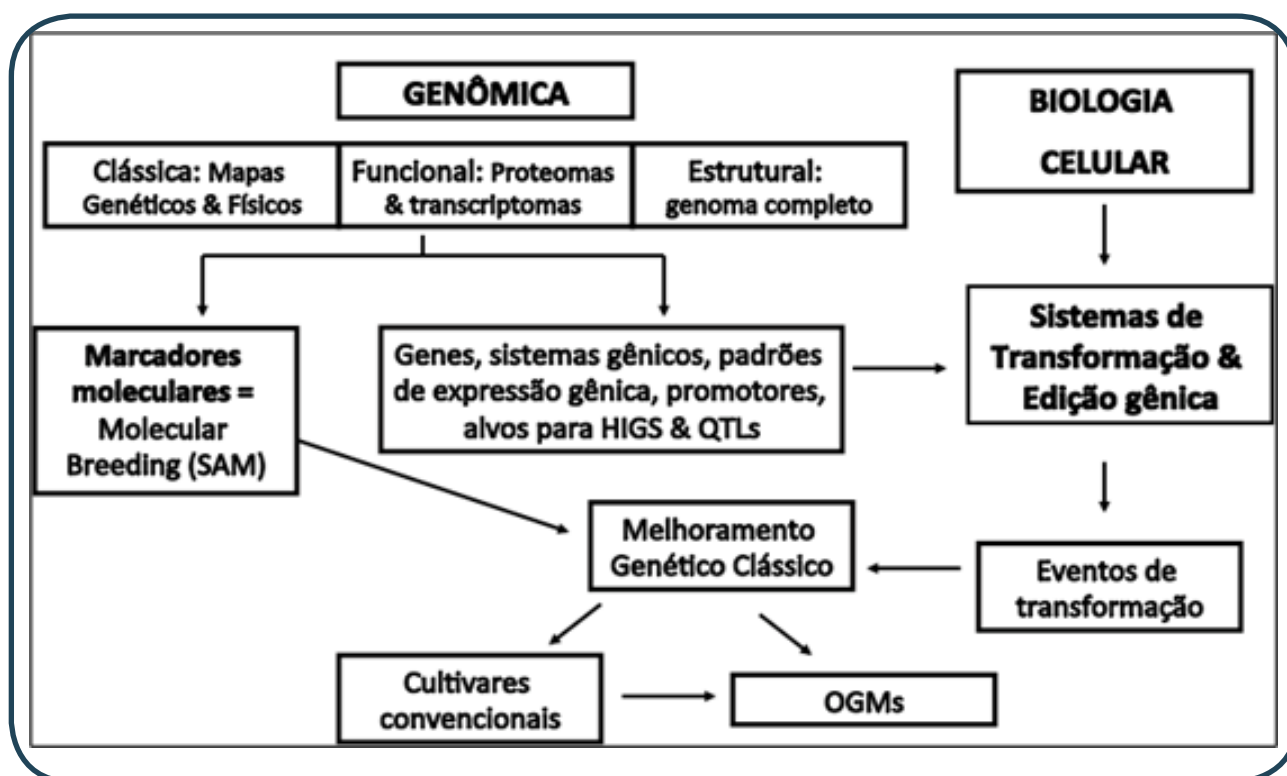


Figura 5. Diagrama das interconexões de diferentes áreas de conhecimento associadas com a seleção assistida por marcadores moleculares (SAM) dentro dos programas de melhoramento genético de nova geração.

- *OGM- Organismos Geneticamente modificados;
- *HIGS- Silenciamento Gênico Induzido pelo Hospedeiro
- *QTLs- Quantitative Trait Loci

3. Marcadores Genéticos

Marcadores genéticos podem ser definidos por diferenças em características fenotípicas bem como qualquer diferença estrutural em segmentos de DNA (codantes ou não), RNA ou proteína (produto gênico) que revelam (direta ou indiretamente) diferenças genéticas entre indivíduos (= **polimorfismos**). Para funcionarem como marcadores, estas diferenças têm que estar em associação (ligação) genética estreita com os caracteres de interesse. Os marcadores moleculares de DNA são os mais abundantes e mais estáveis quimicamente, facilitando os processos de análise (Jones et al., 1997). Devido a essas propriedades, grande parte dos demais marcadores (baseados em RNA ou proteínas) são prioritariamente convertidos para marcadores de DNA. Em geral quanto mais próximo do gene de interesse maior a probabilidade do marcador de DNA de permanecer ligado ao gene. Essa é base teórica que permitiu o desenvolvimento dos primeiros mapas genéticos como originalmente elucidado por Stutervant (1921).

4. Propriedades & tipos de Marcadores Genéticos

Três tipos de marcadores **genéticos** têm sido utilizados, quando disponíveis, na seleção assistida: os morfológicos, os bioquímicos e os moleculares. **Marcadores morfológicos:** resultam da análise direta e simples de diferenças (polimorfismos) no fenótipo (= morfologia) entre indivíduos. **Marcadores bioquímicos:** são obtidos via técnicas de detecção de polimorfismos ao nível de proteína (geralmente enzima) ou padrão de proteínas. **Marcadores moleculares:** são obtidos via técnicas de detecção de polimorfismos ao nível de DNA e RNA (Jones et al., 1997).

4.1. Marcadores morfológicos ou clássicos

Desde Mendel até meados da década de 1960, os marcadores predominantemente utilizados em melhoramento eram associados a caracteres morfológicos. Marcadores morfológicos são aqueles controlados por características fenotípicas de fácil identificação associadas e/ou ligadas com outra característica de interesse (Rick & Yoder, 1988). Os marcadores morfológicos são caracteres visíveis a olho nu (porte da planta, deficiência de clorofila, cor de flores etc.) ou em lupa (formato, tamanho e cor de folhas, frutos, flores e outras partes da planta). No caso do tomateiro os caracteres fruto com ombro verde e arquitetura de folha (folha ereta) podem ser utilizados como marcadores um do outro, ou com outros caracteres localizados no início do cromossomo 10 porque apresentam estreita co-segregação (González-Arcos et al., 2019). De fato, os primeiros mapas genéticos se constituíam exclusivamente em marcadores morfológicos e citológicos. As desvantagens dos marcadores morfológicos incluem o seu pequeno número e devido ao fato de serem identificados, em sua maioria, em diferentes estádios de desenvolvimento da planta. Por sua vez, os marcadores bioquímicos ou fragmentos de DNA, podem ser utilizados a partir de amostras de células ou de tecidos de todas as partes da planta (folhas, embriões, cotilédones, pólen, sementes etc.) e também em qualquer estágio de desenvolvimento da planta. No tomateiro, marcadores morfológicos foram inicialmente mapeados empregando estoques cromossômicos com aberrações tais como aneuploidias, translocações, inserções/deleções e inversões (Rick & Yoder, 1988).

4.2. Marcadores bioquímicos

Os marcadores bioquímicos são enzimas com função similar, mas que diferem em sua estrutura e peso molecular, permitindo sua identificação geralmente por eletroforese (Markert & Moller, 1959). A análise dessas **isoenzimas** é realizada em de três etapas: extração das enzimas do tecido vegetal, separação por eletroforese e coloração por métodos histoquímicos. Estes marcadores proporcionam grande facilidade, pois a obtenção de produtos é evidenciada pela simples observação do gel (Alfenas, 1998). São marcadores predominantemente codominantes (ver definição abaixo), possibi-

tando identificação de variantes alélicas para um mesmo gene. Além disso, a técnica de isoenzimas é de baixo custo, fácil e rápida. Isoenzimas tem contribuído muito para caracterização de cultivares de uma gama de espécies culturas (Alfenas, 1998). No tomateiro, o locus da fosfatase ácida 1 (*Aps-1*) foi a primeira isoenzima empregada como marcador para o gene *Mi-1.2* (**Figura 6**) que confere resistência aos nematoídes-das-galhas (Rick & Fobes, 1974; Medina-Filho, 1980). A proteína fosfatase ácida 1 foi purificada e sequenciada parcialmente. A sequência parcial dos aminoácidos foi utilizada para o desenho de iniciadores de PCR e identificação do gene em tomateiro (Williamson & Colwell, 1991). As principais desvantagens desses marcadores são o relativo baixo número dos sistemas enzimáticos polimórficos além de serem fortemente influenciados por fatores ambientais ou pela fase de desenvolvimento da planta (Alfenas, 1998). Apesar de todas essas limitações, os marcadores bioquímicos prestaram valiosos serviços no melhoramento de plantas (Alfenas, 1998), incluindo no melhoramento do tomateiro (Medina-Filho, 1980).

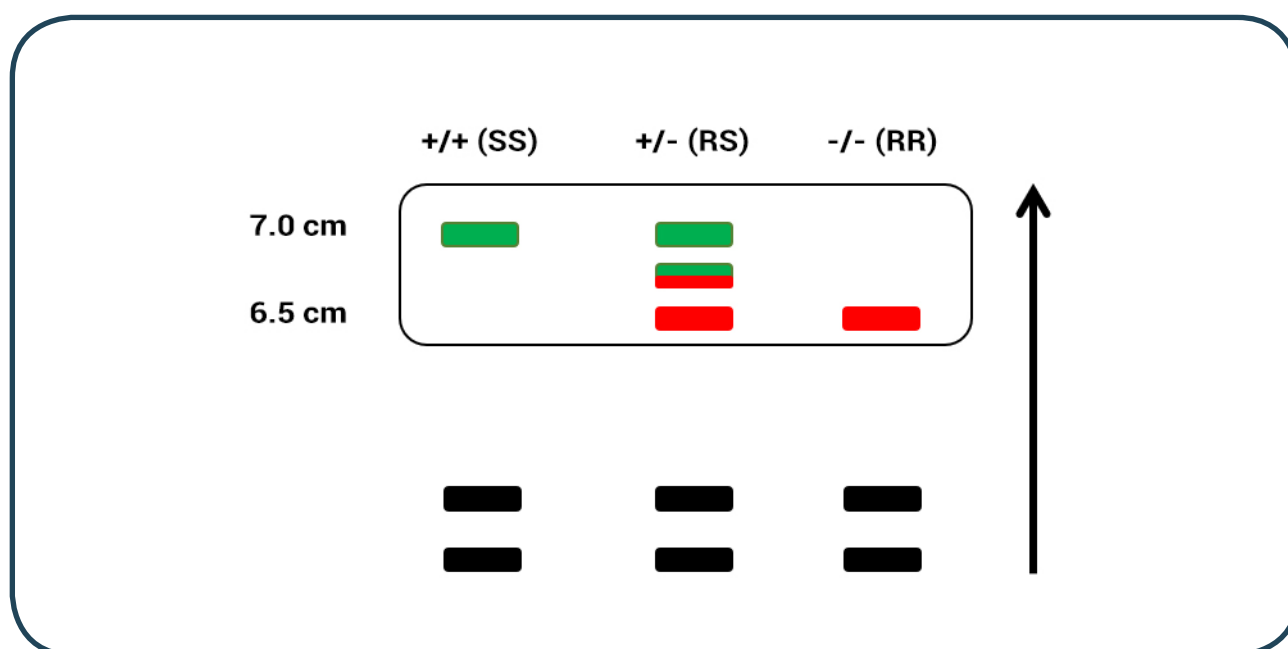


Figura 6. Exemplo do emprego de um marcador bioquímico (isoenzima) no melhoramento genético do tomateiro para resistência aos nematoides-das-galhas (gene *Mi-1.2*). A análise é realizada em de três etapas: extração das enzimas do tecido vegetal, separação por eletroforese e coloração por métodos histoquímicos. As bandas lentas da enzima (vermelhas) codificadas pelos diferentes alelos do gene da fosfatase ácida 1 (*Aps-1*, localizado no cromossomo 6) estão correlacionadas com o genótipo resistente (R) e as bandas rápidas (verdes) aos genótipos suscetíveis (S). Os padrões de bandas para a enzima fosfatase ácida 1 representam indivíduos homozigotos resistentes [banda vermelha (+)], homozigotos suscetíveis [banda verde (-)] e heterozigotos [(+/-)]. No heterozigoto, a enzima ocorre como um dímero (representado por uma banda bicolor) que tem uma migração intermediária. As bandas a 7.0 e 6.5 cm são as derivadas do gene *Aps-1*. As demais bandas (fora da área demarcada) são derivadas de enzimas codificadas por diferentes loci e não estão ligadas com o fator de resistência (gene *Mi-1.2*). Seta indica sentido da migração das bandas em eletroforese. Fonte: Medina-Filho (1980).

4.3. Marcadores Moleculares

Marcadores moleculares são fragmentos de DNA (genes ou regiões não codantes) que são ligados ou codificam uma característica a ser estudada, sinalizando a presença/ausência do gene alvo na planta sob análise. Esta ligação é genética, ou seja, o marcador e o gene estão localizados no genoma de tal maneira que são transferidos predominantemente juntos aos gametas e as progênies (ou seja, co-segregam) (Jones et al., 1997). Isto ocorre porque estão fisicamente ligados no cromossomo e/ou porque não ocorre recombinação frequente entre eles. A frequência de recombinação (em %) é expressa com a unidade cM (centi-Morgan), onde 1 cM é igual a 1% de recombinação entre um marcador e um locus interesse. Ou seja, o marcador e gene de interesse vão juntos no mesmo gameta em uma frequência de 99%. Desta forma, quanto mais forte for esta ligação melhor o marcador. Neste cenário, o marcador molecular perfeito seria aquele derivado da informação advinda do próprio gene (= **marcador funcional**). As modernas tecnologias de análise de marcadores moleculares viabilizam a caracterização genética de um grande número de genótipos através de procedimentos relativamente simples e rápidos. Como consequência, a seleção de indivíduos em programas de melhoramento é realizada de forma mais precisa, determinando a formação de populações segregantes com alta frequência de genótipos superiores (Bered et al., 1997). Marcadores de DNA foram inicialmente utilizados no melhoramento de plantas no início da década de 1980 (Soller & Beckmann, 1983). Ao contrário dos marcadores morfológicos e bioquímicos, os marcadores moleculares são praticamente ilimitados em número, não sendo afetados por fatores ambientais e/ou pela fase de desenvolvimento da planta. Além disso, vários sistemas de marcadores moleculares apresentam fácil detecção e análise e se comportarem como “caracteres” de herança simples (Jones et al., 1997). No processo de transferência de alelos de resistência, os marcadores moleculares podem ser uma ferramenta bastante útil. Esses marcadores, se fortemente ligados aos alelos de resistência, podem ser usados em SAM, sendo imprescindíveis na seleção assistida durante o processo de piramidização de alelos de resistência para um mesmo patógeno ou grupo de patógenos. Essa estratégia vem sendo considerada como uma forma de desenvolver cultivares com resistência duradoura e de amplo espectro.

Tipos de polimorfismos de DNA

Os marcadores moleculares se baseiam na presença de polimorfismos que possam ser utilizados para distinguir organismos com diferenças de interesse. A capacidade de identificar e de detectar tais polimorfismos depende da magnitude dessas diferenças e da precisão da técnica de detecção. Polimorfismos são diferenças entre alelos ou regiões genômicas causadas por mutações do tipo deleção/inserção de mais de uma base (InDels), deleções, inserções ou trocas de uma única base (SNPs), repetições ou modificações epigenéticas no DNA (metilação). No tomateiro, os SNPs representam a maior classe de polimorfismos, sendo o maior número encontrado nas espécies selvagens e preferencialmente localizados em regiões intergênicas (Gupta et al., 2020).

Marcadores dominantes e codominantes

Marcadores dominantes são detectados em apenas um dos alelos e não são visíveis em indivíduos com o alelo contrastante. O marcador codominante é visível e distinto en-

tre os dois alelos, sendo o marcador para o heterozigoto facilmente reconhecido (Jones et al., 1997). O marcador dominante dificulta a análise pois a não visualização do alelo pode ser resultado de alguma falha nos procedimentos laboratoriais e não necessariamente de origem genética. Dessa forma, os marcadores dominantes necessitam de um número maior de controles e repetições do que os marcadores codominantes. A origem bem como os principais contrastes entre os marcadores dominantes e codominantes detectados em ensaios de PCR estão ilustrados na **Figura 7**. A presença de InDels entre os sítios de anelamento dos primers usualmente definem marcadores codominantes, enquanto os SNPs usualmente definem marcadores dominantes. No entanto, quando a presença de um SNPs resulta na formação de um novo sítio de clivagem para uma enzima de restrição, esta informação pode ser convertida na geração de marcadores codominantes.

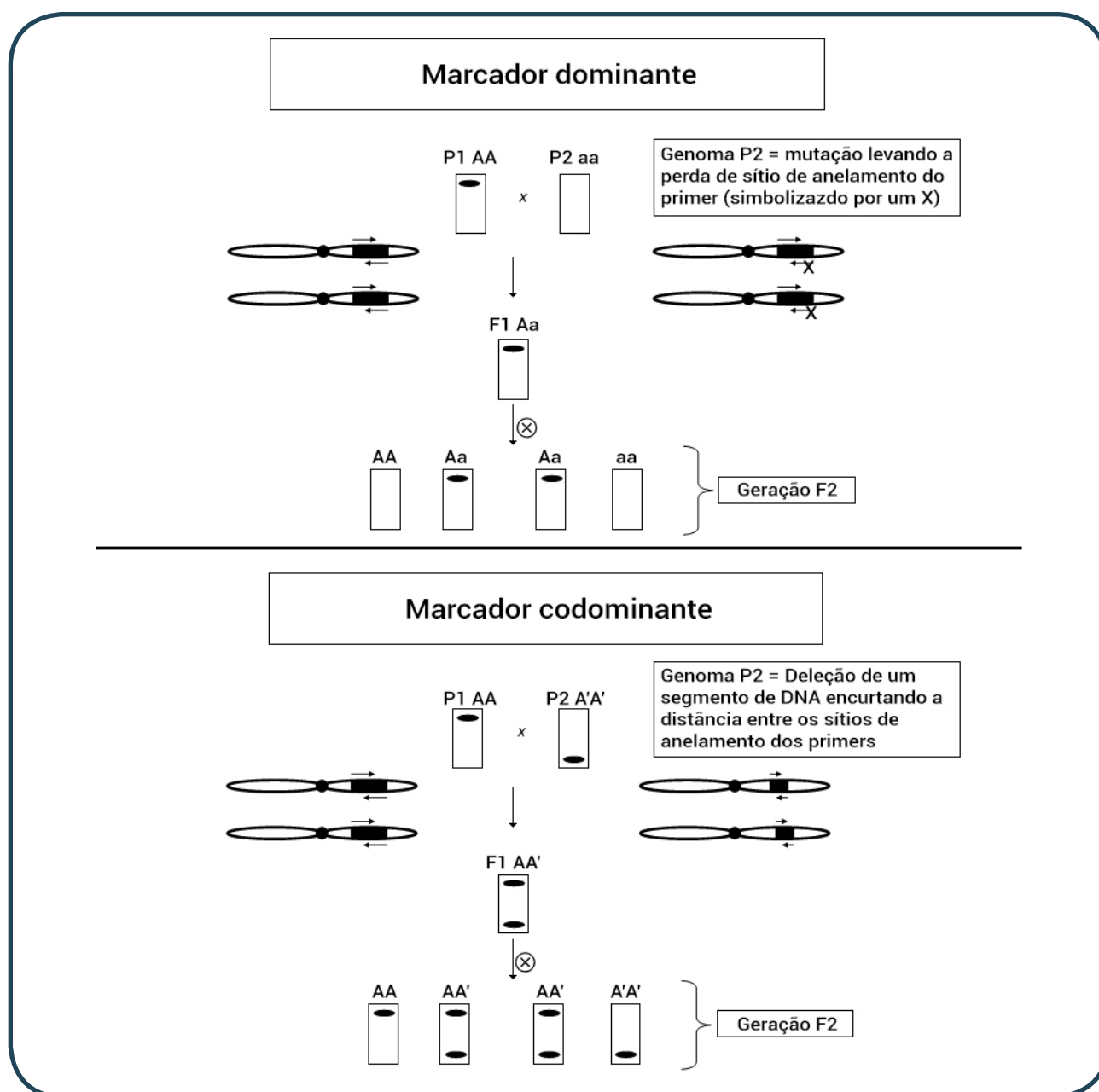


Figura 7. Diagramas contrastando as características peculiares dos marcadores moleculares dominantes e codominantes em ensaios de PCR. No painel superior o local de anelamento dos primers é perdido no indivíduo aa e a região delimitada pelos primers não é amplificada. No painel inferior a deleção no fragmento interno em A'A' gera amplicons menores, mas ainda detectáveis.

5. Principais Categorias e Tipos de Sistemas de Marcadores

No presente capítulo, os principais sistemas de marcadores são agrupados nas seguintes categorias: **(1)** Marcadores moleculares baseados em hibridação; **(2)** Marcadores baseados em PCR com primers arbitrários; **(3)** Marcadores baseados em PCR com primers não-arbitrários; **(4)** Marcadores baseados em PCR & análise de restrição; **(5)** Marcadores baseados em sequenciamento de amplicons polimórficos; **(6)** Marcadores derivados de NGS; **(7)** Marcadores baseados em informação de sequência & PCR em tempo real e **(8)** Marcadores derivados de transcritomas.

5.1. Sistemas de Marcadores Moleculares Baseados em Hibridação

Estes sistemas de marcadores moleculares detectam diferenças de fragmentos de DNA obtidos após corte com enzimas de restrição e hibridação com sondas marcadas (radioativas ou “frias”). **RFLP (Restriction Fragment Length Polymorphism)** – É o principal representante deste grupo de marcadores. Foi o primeiro sistema de marcadores de DNA utilizado no melhoramento de plantas (Helentjaris et al., 1986), sendo os marcadores moleculares utilizados para gerar os primeiros mapas genéticos ultradensos do tomateiro (Tanksley et al., 1992). Nessa técnica, o DNA total de um indivíduo é inicialmente isolado e clivado com enzimas de restrição. Os fragmentos obtidos são separados por eletroforese e transferidos para uma membrana de celulose ou náilon. Em seguida, fragmentos específicos podem ser detectados pela incubação da membrana com uma sonda (uma sequência de DNA marcada). Essa sonda irá, por complementaridade entre as bases nitrogenadas, parear com um ou mais dos fragmentos contidos na membrana. As variações nos nucleotídeos do DNA devido à mutação, inserção/deleção e inversão podem ser detectadas se essas ocorrerem em uma região do genoma que define um sítio de corte de enzimas de restrição. Se o DNA de plantas diferindo em um ou vários desses nucleotídeos forem expostos a essas enzimas, diferentes padrões de fragmentos ou fragmentos de diferentes tamanhos, portanto polimórficos, são gerados e podem ser identificados e clonados. Tais fragmentos são denominados RFLPs e foram inicialmente desenvolvidos por Botstein et al. (1980). Os RFLPs são loci no DNA que podem ser identificados e mapeados. **Minissatélites (VNTRs = Variable Number of Tandem Repeats)** – são unidades de 10 a 100 pb repetidas em tandem, flanqueadas por sítios conservados de endonucleases de restrição. Originalmente, esta metodologia utiliza os mesmos princípios da técnica de RFLP que abrange técnicas de restrição do DNA e uso de sondas para hibridização, sendo extremamente laboriosa. Contudo, esses marcadores podem ser analisados por PCR, de forma mais rápida e prática, uma vez que, as repetições são conservadas no genoma de uma mesma espécie. Muitos dos minissatélites são altamente polimórficos, produzindo um grande número de bandas. Por estarem espalhadas por todo o genoma e apresentarem um número variável de repetições em diferentes indivíduos em rela-

ção a uma mesma região cromossômica (locus), os minissatélites simultaneamente proporcionam um conjunto de marcadores genéticos que se constitui no que tem sido denominado de impressões digitais de DNA, consequentemente, indivíduo-específicos. Para a obtenção do padrão de bandas utiliza-se o mesmo procedimento utilizado para o RFLP, com exceção de que a sonda contém repetições de sequência conhecida. As desvantagens são relacionadas à dominância dos marcadores e ao fato de a técnica ser mais trabalhosa, demorada e cara (Ferreira & Grattapaglia, 1998).

5.2. Marcadores Baseados em PCR com Primers Arbitrários

RAPD (Random Amplified Polymorphic DNA)

No final da década de 1980, foi desenvolvido um novo tipo de sistema de marcadores moleculares denominado de RAPD (Welsh & McClelland, 1990; Williams et al., 1990) ou de **Arbitrarily primed PCR** = AP-PCR (Caetano-Anollés et al., 1991). O RAPD é uma técnica que utiliza a reação de PCR para detectar fragmentos específicos de DNA (Williams et al., 1990). O sistema RAPD fez parte da primeira geração de marcadores obtidos pela utilização de estratégias mais simples via PCR. O RAPD é uma ferramenta bastante útil quando não existem informações sobre os genes a serem amplificados ou sobre o genoma da espécie em estudo. Ao contrário do PCR convencional, o RAPD utiliza apenas um ‘primer’ curto (de até dez nucleotídeos) que, devido ao seu pequeno tamanho, pode parear em diversos pontos do genoma. Caso duas cópias desse ‘primer’ se liguem às fitas opostas do DNA a região flanqueada pelos ‘primers’ pode ser amplificada desde que separada por uma distância amplificável via PCR (entre 200 e cerca de 2.000 pares). O RAPD não exige o conhecimento prévio da sequência que está sendo amplificada, ao contrário do PCR convencional. Logo, um mesmo conjunto de ‘primers’ pode ser usado para amplificar o DNA dos mais diversos tipos de organismos. As principais vantagens da técnica são a facilidade e a rapidez para obtenção dos marcadores, a necessidade de quantidades mínimas de DNA e a universalização das análises, sendo muito empregado no desenvolvimento de marcadores ligados a genes de resistência a doenças como por exemplo no feijoeiro (Kelly & Miklas, 1998) e no tomateiro (Boiteux et al., 2016). A baixa reprodutibilidade tem sido apontada como a grande limitação deste sistema de marcadores. Na verdade, a técnica requer certa experiência com procedimentos moleculares, cuidado com o preparo das reações e rigor na leitura e análise dos fragmentos visualizados no gel. A quantidade e qualidade do DNA, a concentração de íons magnésio e da enzima *Taq* polimerase são alguns dos aspectos que devem ser considerados nos ensaios de RAPD. É importante que o DNA esteja padronizado e não muito concentrado, livre de contaminações de polissacarídeos, proteínas e compostos fenólicos, pois estas substâncias inibem ou diminuem a ação da *Taq* polimerase (Lopes et al., 2012). Em plantas, os RAPDs têm facilitado a condução de estudos em genética e melhoramento, até então, considerados inexequíveis com as técnicas tradicionais. Os polimorfismos detectados via RAPD podem ser InDels ou SNPs. Uma diferença entre o DNA de duas plantas que ocorra na região de anelamento

do *primer* é identificada pela ausência da referida banda em uma delas e presença da banda na outra. No caso de indivíduos heterozigotos, estes produzem as mesmas bandas que os homozigotos, pois os marcadores rápidos são majoritariamente dominantes.

5.3. Marcadores Baseados em PCR com Primers Não-Arbitrários

Os iniciadores da PCR podem ser sequências conhecidas que se deseja amplificar (exemplo: um gene de resistência) ou sequências conservadas que flanqueiam os microssatélites (que são segmentos genômicos repetitivos dispersos ao longo de todo o genoma). Esses marcadores podem também ser derivados de sequências de genes conhecidos contendo regiões genômicas ou domínios conservados (em termos evolucionários) mesmo entre plantas hospedeiras sem estreitas relações filogenéticas. Essa informação pode ser empregada no desenho de primers (= **PCR heterólogo**), tendo como objetivo a descoberta de novos alelos (análogos ou homólogos) em uma nova hospedeira. Outro sistema de marcadores nessa classe é o uso de primers cobrindo regiões conservadas da transcriptase reversa de **retrotransposons**. Esses elementos móveis são interessantes para a potencial geração de marcadores moleculares uma vez que estão dispersos ao longo de todo o genoma de plantas. **Marcadores SSR (Simple Sequence Repeats) ou Microssatélites** – O genoma dos eucariotos apresenta diferentes classes de sequências repetidas. Tais repetições podem ser classificadas de acordo com a sua extensão em: satélites, minissatélites e microssatélites. Marcadores microssatélites ou sequências simples repetidas (SSR) têm sido preferidos a outros tipos de marcadores, pois utilizam a agilidade da técnica de PCR, são codominantes e estão espalhados no genoma em uma frequência alta (Akkaya et al., 1992). Essas sequências são constituídas de repetições de nucleotídeos que ocorrem naturalmente no genoma. Devido a erros que podem ocorrer durante a replicação do DNA, diferentes indivíduos de uma mesma espécie podem apresentar um número variado de repetições dentro de um mesmo microssatélite (Akkaya et al., 1992). Para que o microssatélite seja útil como marcador, ele deve ser inicialmente identificado, clonado, sequenciado e amplificado a partir de ‘primers’ que o flanqueiem. Alternativamente podem ser identificados diretamente no genoma, caso esse esteja disponível. Após o desenvolvimento do marcador a rotina passa a ser a PCR e a separação dos produtos por eletroforese. Na maioria das vezes, a eletroforese deve ser feita em gel de poliacrilamida devido à pequena diferença de tamanho entre os fragmentos alélicos polimórficos. Os elementos repetitivos são formados por arranjos de repetições em “tandem”, de dois a seis nucleotídeos de comprimento e são os loci mais polimórficos dos genomas (Milach, 1998). Por exemplo, a sequência GACAGACAGACAGACA, também representada $(GACA)_4$, é um microssatélite. O polimorfismo desses marcadores baseia-se na variação do número dos elementos repetidos, provavelmente devido aos erros da DNA polimerase durante o processo de replicação e reparo da molécula de DNA. Em genomas de eucariotos estas sequências estão distribuídas ao acaso e constituem uma das classes mais polimórficas de marcadores moleculares disponíveis, devido à sua alta taxa de mutação. Ao usar-se primers específicos que flanqueiam o microssatélite para sua amplificação,

têm-se diferenças entre indivíduos de acordo com o número de repetições da sequência, ou seja, bandas de tamanhos diferentes serão produzidas. Os microssatélites são marcadores codominantes e cobrem bem o genoma, entretanto, para o uso rotineiro existe a necessidade de primeiro amplificar uma região, posteriormente sequenciar e em terceiro lugar, sintetizar os iniciadores específicos para cada locus. Uma vez feito isto, o locus marcador pode ser utilizado indefinidamente naquela espécie. Desta forma, existe um custo elevado e trabalho no início, mas o custo subsequente é baixo e a simplicidade a posteriori, é muito grande.

5.4. Marcadores Baseados em PCR & Análise de Restrição

AFLP (Amplified Fragment Length Polymorphism)

AFLP é uma técnica que combina a clivagem de fragmentos de DNA com enzimas de restrição e amplificação desses fragmentos por PCR (Vos et al., 1995). Nessa técnica, o DNA é clivado com enzimas de restrição, às suas extremidades são ligados adaptadores, os quais servem de sítios de ligação para ‘primers’ numa reação de PCR. Os marcadores AFLP combinam a sensibilidade da detecção por PCR com a ubiquidade dos sítios de restrição no genoma de plantas para busca de polimorfismos (Vos et al., 1995). Assim como o RAPD existem coleções e “kits” contendo estas combinações. No caso do AFLP a estratégia tem uma patente e tais kits são comercializados por empresas de insumos para Biologia Molecular. Para obtenção de marcadores AFLP o primeiro passo é a digestão do DNA por duas enzimas de restrição: uma de corte raro e outra de corte frequente. É fundamental que a digestão do DNA seja completa, pois a digestão parcial pode revelar falsos polimorfismos. A técnica baseia-se na propriedade de certas enzimas de restrição de deixar, após a clivagem do DNA, extremidades coesivas de sequência conhecida. Assim, é possível construir sequências de nucleotídeos de fita dupla que se ligam às extremidades dos fragmentos de restrição, denominadas de adaptadores. Uma vez conhecidas a sequência dos adaptadores e a do sítio de restrição, podem-se construir iniciadores específicos a essas sequências para pré-amplificação dos fragmentos de restrição. Esta etapa consiste na amplificação dos fragmentos agora ligados aos adaptadores através da reação da polimerização em cadeia com o uso de iniciadores, complementares aos adaptadores com uma base a mais. A amplificação final é feita com uma pequena amostra da primeira amplificação. Neste caso são utilizados iniciadores que são compostos de todas as bases dos *primers* da primeira amplificação, mais duas a três bases, dependendo do nível de polimorfismo da espécie ou da população. Existem inúmeras vantagens no uso da técnica de AFLP, como alto número de bandas analisadas em um único gel, pela restrição e amplificação de fragmentos espalhados por todo o genoma, além da considerável estabilidade do perfil de amplicons e, principalmente, por não necessitar de dados de sequenciamento prévio da espécie em estudo (Spooner et al., 2005). Apesar de ser um tipo de marcador bastante útil para a realização de *fingerprints* de DNA, principalmente quando existem poucas informações disponíveis a respeito do genoma de interesse, o seu uso no melhoramento de plantas tem sido limitado devido a dificuldades metodológicas inerentes à técnica e ao seu elevado custo.

5.5. Marcadores baseados em sequenciamento de amplicons polimórficos

São marcadores moleculares derivados dos anteriores. Após purificação e sequenciamento do fragmento polimórfico esse pode ser convertido em um marcador mais simples, baseado apenas em PCR e/ou PCR mais análise de restrição. **CAPS (Cleared Amplified Polymorphic Sequence)** – CAPS são fragmentos de DNA amplificados via PCR, utilizando-se de primers específicos (20 a 30 pb), seguido da digestão com endonucleases de restrição. As etapas para obtenção desses marcadores são a extração e amplificação de DNA via PCR, a digestão com enzimas de restrição, eletroforese em géis de agarose e a visualização dos polimorfismos pela coloração com brometo de etídeo sob luz ultravioleta. As principais vantagens são a codominância e a alta reprodutibilidade dos marcadores e a principal desvantagem é a necessidade de conhecimento prévio de sequência de DNA para a construção dos primers. Existem ferramentas online para o desenho de CAPS (visitar https://solgenomics.net/tools/caps_designer/caps_input.pl). **SCAR (Sequence Characterized Amplified Region)** – O sistema RAPD apresenta uma desvantagem que limita, até certo ponto, o seu uso generalizado. A baixa temperatura necessária para a ligação do ‘primer’ ao molde (em torno de 37 °C) permite que muitos amplicons não específicos sejam visualizados. Isso torna o processo bastante dependente das condições de amplificação. Portanto, modificações no termociclador utilizado, na DNA polimerase e em outros reagentes do mix de reação podem alterar o padrão de amplificação. Para contornar essa limitação, marcadores RAPD podem ser transformados em marcadores SCAR. Neste caso, a banda de DNA correspondente ao marcador RAPD é clonada, sequenciada e dois ‘primers’ (mais longos que o original) são sintetizados e utilizados para amplificar o mesmo marcador, só que agora em uma temperatura de anelamento mais elevada. Nessa nova condição, o processo de amplificação é mais estável e específico (Paran & Micheltore, 1993). As etapas principais no desenvolvimento de um SCAR são: **(1)** identificação de um iniciador que confere polimorfismo a dois bulks de DNA com fenótipos contrastantes; **(2)** o isolamento e a clonagem do fragmento amplificado em um vetor (plasmídeo); **(3)** sequenciamento do fragmento isolado; **(4)** desenho dos iniciadores de tamanho maior que os decâmeros e **(5)** o teste ou validação final (Paran & Micheltore, 1993). Até a etapa 3 usa-se de marcadores RAPDs. De posse da sequência, se desenharam os *primers*/iniciadores (etapa 4) com comprimento variável entre 16 e 24 pares de bases. A ideia de um iniciador mais longo surgiu de cálculos feitos sobre o comprimento mínimo para a estabilidade de um *primer* amplificar uma sequência única de um genoma vegetal. Desta forma, espera-se a geração de uma banda única com o uso dos referidos iniciadores. Existem programas de computador que auxiliam a tomada de decisão para a escolha do par de primers específicos mais adequados como, por exemplo, o Primer - Blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). Esses aplicativos computacionais proporcionam valiosas informações comparativas a respeito de diferentes iniciadores que são gerados quando uma determinada sequência de bases é fornecida ao programa.

5.6. Marcadores derivados de NGS

Os avanços tecnológicos e diminuição nos custos tem permitido o sequenciamento de genomas de cultivares/acessos do tomateiro cultivado e de espécies selvagens comumente fontes de alelos de resistência a doenças. Este banco de dados é uma fonte de informações de sequências que pode ser utilizada para a descoberta de novos marcadores moleculares e consequentemente de novos genes. O sequenciamento do tipo **NGS** (Shendure & Ji, 2008) aumenta significativamente a eficiência e reduz os custos em relação aos métodos anteriores. Com base nos avanços dessa técnica, foi possível concluir o projeto completo de sequenciamento do genoma do tomate em 2012. O primeiro modelo sequenciado (no caso do tomateiro a cultivar ‘Heinz 1706’) e a evolução de diferentes hardwares e softwares capazes de armazenar, processar e analisar quantidades significativas de dados, facilitou o resequenciamento de diferentes genomas e transcritomas. Aplicando esta ferramenta, também foi possível descobrir inúmeras variantes através do genoma e o desenvolvimento de genotipagem de larga-escala. Os genomas derivados de resequenciamento estão disponíveis em solgenomics.net. Após a análise em diversos programas de alinhamento disponíveis pode-se escolher regiões com SNPs e InDels para o desenho de marcadores. À medida que os tempos de sequenciamento e os custos são reduzidos, novas estratégias de genotipagem estão agora disponíveis. Com o resequenciamento do genoma foi possível identificar dezenas de milhares de variantes (exemplo SNPs) ao longo do genoma do tomate (Hamilton et al., 2012). Com o desenvolvimento paralelo de plataformas de genotipagem (Gupta et al., 2020), que facilitam a caracterização de populações a partir de marcadores selecionados, foi possível construir novos mapas genéticos de alta densidade (Sim et al., 2012). Sua aplicação permite localizar com precisão no genoma uma característica de interesse facilitando o processo de seleção de genes candidatos. Outras estratégias de genotipagem (baseadas em sequenciamento de alto desempenho) estão sendo usadas atualmente. O objetivo da estratégia é identificar polimorfismos (SNPs) e coletar simultaneamente informações genômicas de uma população de interesse. Para isso, utiliza-se o sequenciamento de fragmentos de DNA específicos com representação genoma mais ou menos total de indivíduos (ou grupos de indivíduos) de uma população. Mais recentemente, a técnica **resequenciamento (Re-seq)** tem sido empregada usada para encontrar um grande número de marcadores genéticos. Além disso, tecnologias baseadas na estratégia de **digestion-based simplified genome sequencing** também têm sido usadas, incluindo o **RAD-seq = restriction site associated DNA sequencing** (Baird et al., 2008), **genotyping-by-sequencing (GBS-seq)** ou genotipagem por sequenciamento (Elshire et al., 2011; Kim et al., 2016) e a **SLAF-seq = specific-locus amplified fragment sequencing** (Sun et al., 2013). A técnica SLAF-seq tem a vantagem de evitar sequências de genoma repetitivas se tornando altamente precisa, de baixo custo e rápidas. Atualmente, essas metodologias têm sido aplicadas em genotipagem de larga escala e para várias espécies, com ou sem genomas de referência. Essas metodologias são usadas

na construção de mapas genéticos de alta densidade e empregadas para mapear diretamente características/genes individuais. Como exemplos no tomateiro, temos a aplicação da metodologia RAD-seq para o reposicionamento de fatores quantitativos relacionados à resistência ao oomiceto *Phytophthora infestans* (Chen et al., 2014) e o emprego da metodologia SLAF-seq para o mapeamento de um gene relacionado à resistência ao fungo *Cladosporium fulvum* (Zhao et al., 2016). Uma abordagem interessante tem sido o emprego da metodologia de **BSA-seq** (Win et al., 2017) para mapear genes de resistência (exemplo: gene recessivo *ty-5* que confere resistência a begomovírus; Wang et al., 2018). **BSA-seq** combina modernas técnicas de sequenciamento com o método de **Bulked Segregant analysis – BSA** (Michelmore et al., 1991) para rapidamente ‘aterriçar’ em genes controlando tanto características qualitativas quanto quantitativas.

5.7. Marcadores baseados em informação de sequência & PCR em tempo real

Os polimorfismos do tipo SNPs, InDels e mesmo em modificações induzidas por metilação podem ser identificados na técnica de **High Resolution Melting (HRM)**. Na análise por HRM os polimorfismos são revelados após amplificação, pelo padrão de desnaturação dos amplicons. Os marcadores do **tipo KASP (= Kompetitive Allele Specific PCR)** se baseiam na marcação dos alelos de cada parental com um fluorômetro específico (Semagn et al., 2014).

5.8. Marcadores derivados de transcritomas

Transcritoma é o conjunto completo de transcrições de uma célula e sua quantificação, associada a um determinado estado de desenvolvimento e condição ambiental. A tecnologia de sequenciamento de alto desempenho permitiu gerar ferramentas como **RNA-seq** (Wang et al., 2009) capazes de mapear e quantificar transcrições com níveis muito altos de eficiência (tempo, custo). Usando RNA-seq, juntamente com um design experimental adequado, transcrições expressas diferencialmente podem ser identificadas contra um estímulo ambiental e em determinado tecido vegetal, de modo que esta ferramenta pode ser usada especialmente para selecionar genes candidatos por função, com o benefício de que, ao mesmo tempo, se um genoma de referência estiver disponível, esses genes também são mapeados. É o caso da seleção de genes candidatos para o controle do formato do fruto de tomate (Sun et al., 2013) e a identificação do gene *I-7* que confere resistência ao fungo *Fusarium oxysporum* f. sp. *lycopersici* raças 1, 2 e 3 (González-Cendales et al., 2016).

6. Construção e aplicação dos mapas genéticos (linkage maps) com alta densidade de marcadores moleculares no isolamento de Genes e identificação de QTLs

Diversos genes de resistência a doenças têm sido isolados, utilizando as estratégias de mapeamento genético/físico de alta resolução (**map-based cloning ou positional cloning**). Esse tipo de trabalho tem o objetivo de encontrar marcadores moleculares próximos/ligados ao gene de interesse, saturando a região onde o gene se localiza, com vários marcadores, o que vem sendo acelerado/facilitado em algumas espécies cujos genomas completos já estão disponíveis. Os **mapas de ligação (linkage maps)** são construídos com base em análises da segregação dos marcadores em uma população segregante, geralmente na geração F_2 . Na construção destes mapas, quando marcadores polimórficos são encontrados, eles devem ser validados em toda a população de mapeamento, que é a **genotipagem** da população. Portanto, o DNA de cada indivíduo da população de mapeamento deve ser extraído. As taxas de segregação esperadas para marcadores dominantes e codominantes para uma população F_2 são, respectivamente, 3:1 e 1:2:1 (Collard et al., 2005). Desvios significativos entre taxas esperadas e taxas observadas, podem ser analisados usando o teste qui-quadrado. Geralmente, os marcadores segregarão em um modelo Mendeliano, embora taxas de segregação distorcidas possam ser encontradas, especialmente em cruzamentos interespecíficos envolvendo espécies cultivadas e selvagens (Sayed et al., 2002; Xu et al., 1997). A última etapa para a construção de um mapa de ligação envolve a análise dos dados obtidos de cada marcador de DNA de cada um dos indivíduos da população e a análise da ligação utilizando softwares com essa função, sendo os mais usados: Map-maker/EXP (Lander et al., 1987), MapManager QTX (Manly et al., 2001) e JoinMap (Stam, 1993). Algumas análises mais simples podem ser feitas de forma manual, dividindo o número de indivíduos recombinantes para os fatores analisados pelo total de indivíduos da população, contudo, para construir mapas com grande número de marcadores, é recomendado o uso de softwares. A acurácia para medir a distância genética e determinar a ordem dos marcadores está diretamente relacionada com o tamanho da população de mapeamento. Populações segregantes ideais para construção de mapas de ligação devem possuir no mínimo 50 indivíduos (Young, 1996). A distância no mapa de ligação é medida em termos de frequência da recombinação entre marcadores genéticos (Paterson, 1996), essa distância é medida em centiMorgans (cM), pois a frequência de recombinação e a frequência de *crossing-over* não são linearmente relacionadas (Paterson, 1996). Quanto maior a distância entre e dos marcadores com o fator estudado, maior a chance de ocorrer recombinação durante a meiose. Quando as distâncias mapeadas são pequenas (<10cM), a distância mapeada se iguala à frequência de recombinação. Entretanto, esta relação não se aplica a distâncias mapeadas maiores que 10 cM (Paterson, 1996). Com o desenvolvimento de mapas genéticos superdensos, vários marcadores moleculares em estreita ligação ou mesmo dentro de alguns desses genes (= marcadores funcionais) já estão disponíveis para ser utilizados em seleção assisti-

da. Além disso, esses mapas de alta densidade são fundamentais para localizar regiões do genoma que possuem genes controlando a expressão de características complexas (poligênicas) são denominadas de **QTLs (= quantitative trait loci)** (Tanksley, 1993; Tanksley & Fulton, 2007). Os QTLs associados com uma dada característica quantitativa podem estar localizados em diferentes cromossomos e apresentar diferentes contribuições na expressão do fenótipo em estudo. Vários QTLs associados com diferentes caracteres de efeito quantitativo já foram identificados estando associados com a chamada resistência horizontal a diferentes doenças (Young, 1996).

7. Aplicações das ferramentas de biologia avançada e de análise genômica disponíveis para o melhoramento do tomateiro

No tomateiro, esses grandes avanços da genética molecular se concentraram no desenvolvimento e uso de marcadores moleculares, a fim de facilitar o mapeamento de genes de interesse e sua transferência para linhagens elite (Eshed & Zamir, 1995; Saliba-Colombani et al., 2000). O tomate foi uma das primeiras espécies cultivadas para as quais marcadores genéticos e mapas foram desenvolvidos para fins de reprodução (Shirasawa & Hirakawa, 2013). Marcadores moleculares baseados em PCR e seleção assistida por marcadores estenderam os limites de seleção fenotípica tradicional (Foolad & Panthee, 2012). Dados moleculares de DNA mitocondrial, RFLPs, microssatélites, isoenzimas e a análise de sequência do gene nuclear *Granule-bound starch synthase* (GBSSI = *waxy*) foram a base de estudos de reavaliação taxonômica do tomateiro que culminaram com a transferência esta espécie bem como espécies correlatas do gênero *Lycopersicon* para o gênero *Solanum* (seção *Lycopersicon*) (Spoonier et al., 2005). Na última década, as ferramentas disponíveis para identificar genes de resistência candidatos revolucionaram esse procedimento. Sobretudo, como comentado, a partir da evolução da eficiência dos métodos de sequenciamento.

7.1. Obtenção dos genomas completos de diversas cultivares do tomateiro e seu potencial uso no melhoramento genético da cultura

Em 2012, o primeiro genoma de referência (da linhagem de ‘Heinz 1706’) foi obtido e disponibilizado após um trabalho coordenado por um consórcio de 14 países que sequenciaram todos os doze cromossomos do tomateiro (The Tomato Genome Sequencing Consortium, 2012). Após isso, foram conduzidos trabalhos de resequenciamento do genoma completo de um grande número de variedades de tomateiro e acessos de espécies silvestres do gênero *Solanum* (seção *Lycopersicon*) (The Tomato

Genome Sequencing Consortium, 2014). Antes disso, diversos genes de resistência a doenças já haviam sido localizados nos diferentes cromossomos via mapeamento genético e físico (= posicional) de alta resolução (Barone et al., 2009). Essas informações criaram novas possibilidades de pesquisa em termos de identificação e localização genética, estudos de função genética, diversidade genética e evolução, não apenas no tomate, mas também em forma comparativa com outras espécies da família Solanaceae (Menda et al., 2013). As informações geradas sobre o genoma do tomate (sequências, genes, mapas físicos e genéticos) bem como sua anotação (International Tomato Annotation Group – ITAG). Dados de transcrito, estruturas de proteínas, previsão de funções de genes e desenvolvimento de vários softwares utilitários para bioinformática são gerenciados e disponíveis no porta da Sol Genomics Network (<http://solgenomics.net/>) (Bombarely et al., 2010) e também no National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov>). Além disso, esses *sites* armazenam informações sobre outras espécies de plantas, permitindo análises genômicas comparativas. O genoma completo de diversas cultivares de polinização aberta de *S. lycopersicum* e de outras espécies selvagens tem sido rapidamente obtido via sequenciamento de nova geração (Lin et al., 2014). As informações moleculares advindas desses projetos têm permitido a síntese de primers de PCR altamente específicos que estão sendo empregados para gerar uma nova classe de marcadores moleculares associados com fenótipos de interesse.

7.2. Identificação de genes de resistência via mapeamento

Como mencionado, os marcadores moleculares ideais são aqueles denominados **marcadores funcionais**, ou seja, derivados dos próprios genes que conferem a característica de interesse. Diversos genes de resistência a doenças têm sido isolados, utilizando as estratégias de mapeamento genético/físico de alta resolução (**map-based cloning** ou **positional cloning**). Esse tipo de trabalho tem como objetivo inicial encontrar marcadores moleculares próximos/ligados ao gene de interesse e como objetivo final identificar o gene responsável pela expressão da característica e utilizar essa informação para gerar um marcador funcional (preferencialmente codominante). Esse trabalho de **clonagem posicional** pressupõe a saturação da região genômica onde o gene se localiza com uma densidade grande de marcadores moleculares. Esse trabalho vem sendo acelerado/facilitado em algumas espécies cujos genomas completos já estão disponíveis. O primeiro gene de resistência a ser isolado em tomateiro e um dos primeiros genes de resistência em plantas foi gene *Pto* que controla resistência a bactéria *Pseudomonas syringae* pv. *tomato* e codifica uma quinase (Martin et al., 1993). O gene *Pto* foi identificado como parte de um *cluster* (agrupamento) de genes que foram isolados via clonagem posicional (**Figura 8**). Outro membro deste *cluster* fisicamente ligado ao *Pto* é o gene *Prf* que é necessário para conferir completa resistência a bactéria (Salmeron et al., 1996). Outro gene ligado aos genes *Pto* e *Prf* é o gene *Fen* que confere sensibilidade ao inseticida Fenthion (Salmeron et al., 1996). Ensaios com este inseticida têm sido empregados com sucesso como um método de seleção indireta para resistência a *P. syringae* pv. *tomato*. Plantas resistentes apresentam uma reação de fito-

toxidez ao inseticida, apresentando pequenas pontuações necróticas após a aplicação do produto. Marcadores codominantes (derivados do gene *Pto*) têm sido empregados em sistemas de seleção assistida, com ilustrado na **Figura 9**.

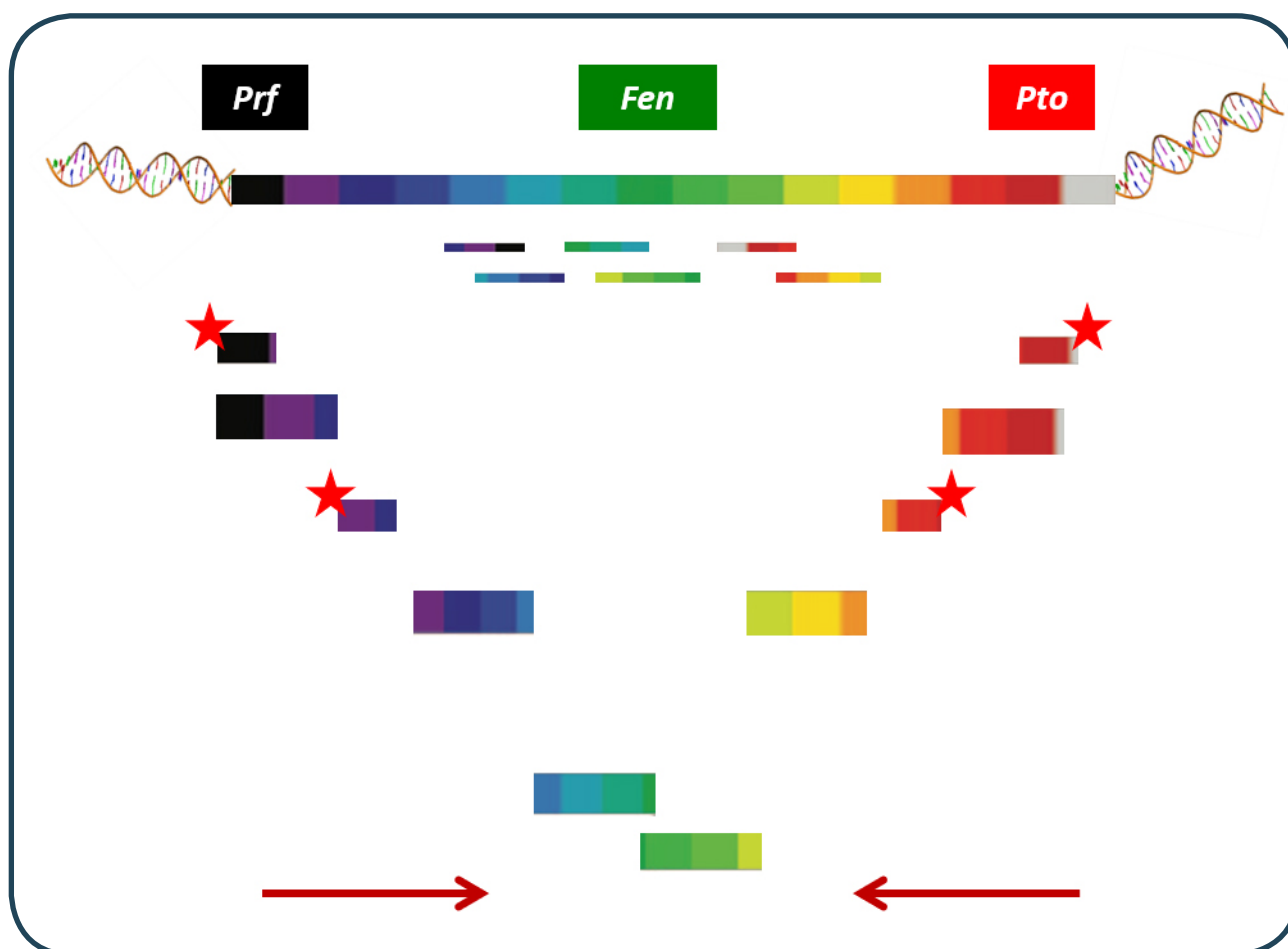


Figura 8. Estratégia de "chromosome walking" no cromossomo 5 do tomateiro: Clonagem posicional do gene de sensibilidade ao inseticida Fenthion (*Fen*) usando marcadores moleculares para as regiões flangeadoras correspondendo aos genes *Prf* & *Pto*, que conferem resistência a bactéria *Pseudomonas tomato* pv. *tomato*. Nesta estratégia, os insertos de clones (DNA clivado e inserido em vetores) contendo os genes *Prf* e *Pto* tem sua sequência definida e são desenhadas novas sondas para sua extremidade. A ligação destas sondas com o fenótipo deve ser checada na população segregante. As novas sondas devem estar mais próximas (= menor número de recombinantes) do gene *Fen* que as sondas iniciais, demonstrando o avanço na direção do gene. Estes ciclos são repetidos até que o número de recombinantes seja zero e um mesmo clone seja selecionado em ambas as direções.

Símbolo das sondas: ★

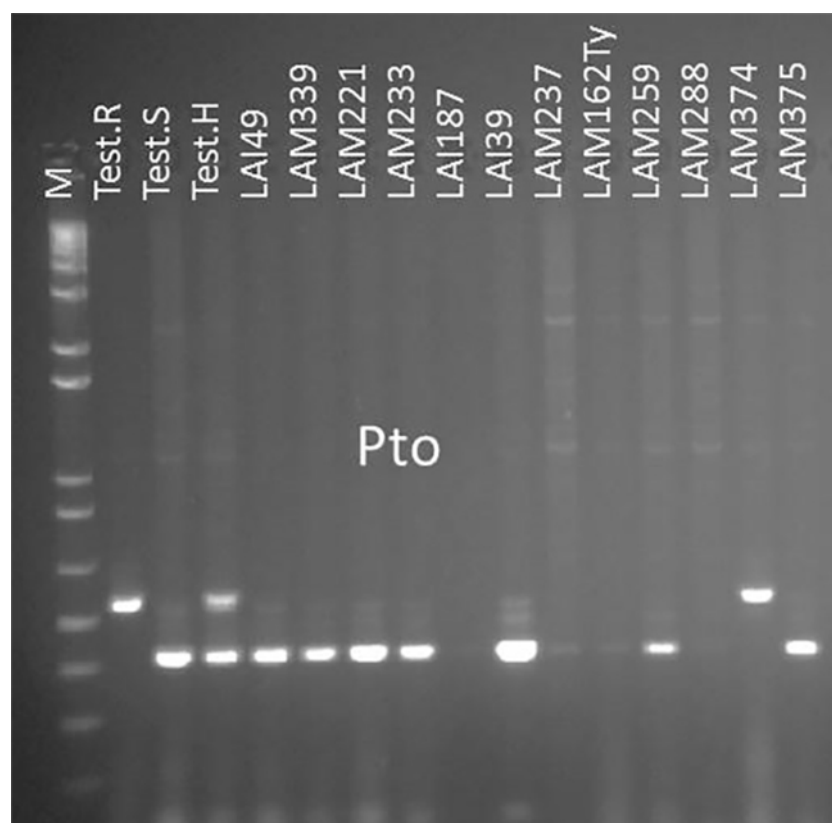


Figura 9. Análise em gel de agarose um marcador molecular codominante ligado ao gene *Pto* que confere resistência a bactéria *Pseudomonas tomato* pv. *tomato*. A testemunha homozigota resistente (Test. R) apresenta uma banda alta, a testemunha homozigota suscetível (Test. S) apresenta uma banda baixa e a testemunha heterozigota para o gene *Pto* (Test. H) apresenta dupla-bandas. A linhagem LAM 374 (parental do híbrido BRS Zamir) apresenta o gene *Pto* em homozigose. Para detalhes metodológicos ver Orsi et al. (2012).

7.3. Desenvolvimento de linhagens de introgressão

As linhagens de introgressão, conhecidas por sua sigla ILs (= *Introgression Lines*) são uma poderosa ferramenta para estudar a localização cromossômica de genes de resistência bem como QTLs (Eshed & Zamir,1995). Cada uma das ILs tem apenas uma única região genômica introgridida e, para todo o resto de genoma, essa linhagem é completamente idêntica a um parental conhecido (**Figura 10**). Esse o conjunto de introgressões individuais cobrem todos os distintos cromossomos. Como resultado, a variação fenotípica observada em uma IL particular pode ser associada com a correspondente região que foi introgridida via monitoramento por marcadores moleculares. Diferentes conjuntos de ILs foram obtidos em tomate, sendo o mais utilizado o conjunto derivado da espécie selvagem *S. pennellii* (Eshed & Zamir,1995). Regiões genômicas apresentando efeitos fenotípicos para diferentes características tais como rendimento, qualidade de fruto e resistência a doenças foram identificados usando essa ferramenta.

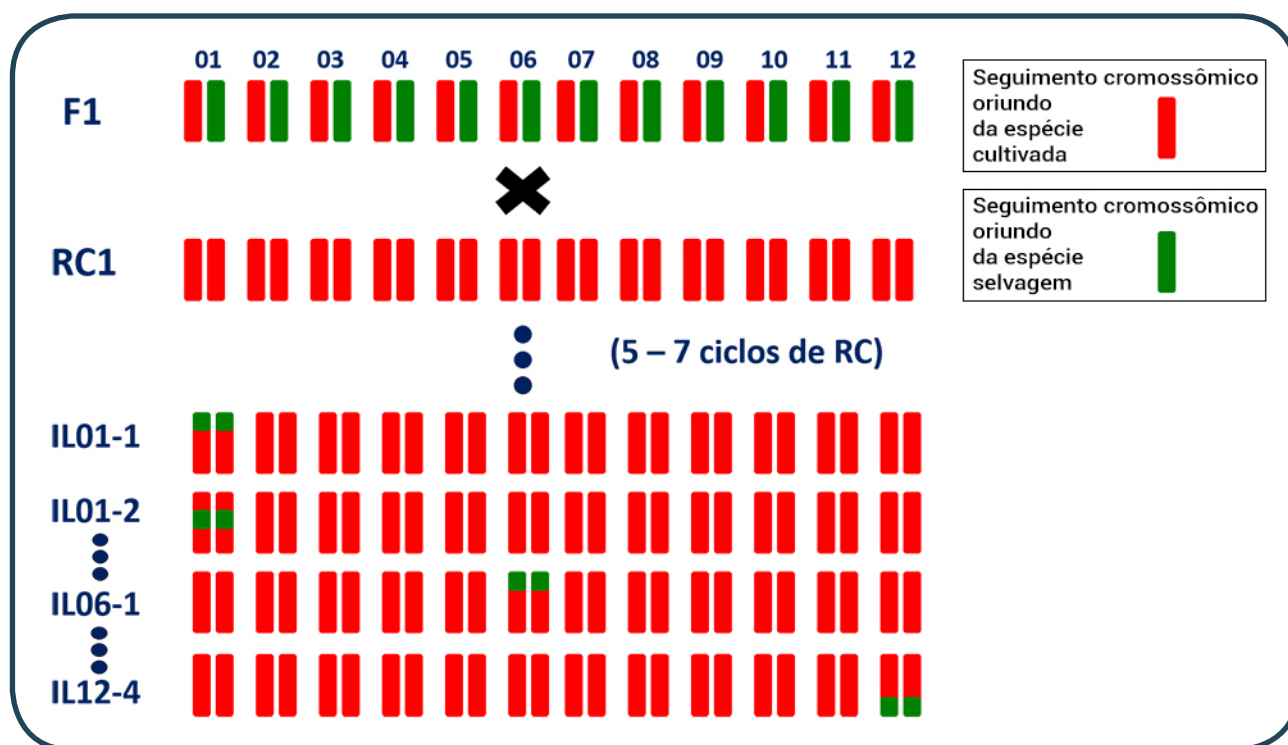


Figura 10. Esquema ilustrando a obtenção de linhagens de introgressão com o auxílio de marcadores moleculares.

7.4. Desenvolvimento de sistemas de marcadores derivados de genes análogos de resistência (RGAs)

A informação inicial sobre a estrutura dos genes de resistência a patógenos que foram clonados permitiu o estabelecimento de métodos mais eficientes e/ou simplificados, baseados em homologia com genes análogos/homólogos (**homology-based methods**), ou na busca por genes candidatos (Tomita et al., 2011). A análise desses genes sugere que as plantas desenvolveram um sistema de defesa que é conservado e que envolve receptores para reconhecimento dos patógenos. Os patógenos por sua vez podem silenciar esses mecanismos pela produção de proteínas efetoras que inibem o sistema imune tornando a planta suscetível. No processo de co-evolução, alguns desses efetores podem ser reconhecidos por uma segunda linha de defesa que envolve genes de resistência que induzem genes de imunidade (effector-triggered immunity – ETI). Esses genes de resistência a doenças em plantas (também denominados de genes R) são classificados em diferentes famílias de acordo com elementos conservados na estrutura das proteínas codificadas. A maioria das proteínas R contém uma região do tipo **nucleotide-binding site (NB)** e um domínio C-terminal do tipo **LRR (= leucine rich repeat)**. Um subconjunto destas famílias gênicas também apresenta um domínio do tipo **coiled-coil (CC)** na região 5' terminal. Os genes R dominantes clonados e estruturalmente caracterizados até o presente momento se enquadram como típicos genes R (Baggs et al., 2017) conhecidos como raça-específicos ou isolado-específicos que funcionam dentro do modelo de interação gene- a-gene. O reconhecimento de fatores de avirulência codificados pelo genoma de uma ampla gama de patógenos e a especifi-

cidade são condicionados pelas regiões C-terminais dos domínios LRR. O domínio NB funciona como um acionador molecular que regula a ativação da cadeia de transdução de sinais que leva ao fenótipo resistente (Baggs et al., 2017). O conhecimento dessas estruturas evolucionariamente conservadas dos genes de resistência pode ser utilizado em ferramentas de alinhamento de genomas de genótipos contrastantes buscando genes candidatos (**Figura 11**). Esses podem ser validados em populações segregantes visando busca de marcadores moleculares ou mesmo o próprio gene de resistência. Iniciadores ou “primers” específicos para genes de resistência conhecidos podem ser desenhados e utilizados para amplificar por PCR genes análogos em cultivares da mesma espécie, diferentes espécies e até diferentes gêneros. A especificidade destes primers depende do número de aminoácidos com múltiplo códon na sequência. O aminoácido Leucina apresenta o maior número de códon, o que força a utilização de uma coleção de primers com sequências variadas nesta posição. Uma alternativa também válida, embora mais cara, é empregar o nucleotídeo sintético Inosina, que forma par com qualquer base (A, C, G ou T). Após o desenho e síntese dos primers utiliza-se o DNA dos genótipos contrastantes como molde na busca de polimorfismos ligados a característica de interesse.

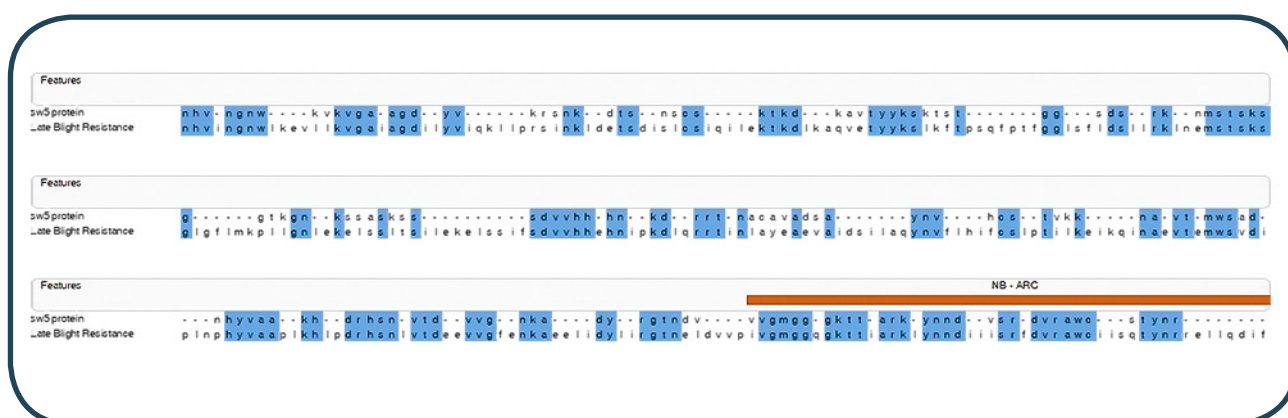


Figura 11. Alinhamento da sequência das proteínas de dois genes de resistência a doenças de tomateiro altamente polimórficos mostrando motivos idênticos (conservados) em azul. Iniciadores de síntese (“primers”) podem ser desenhados nessas regiões e o resultado da amplificação por PCR pode revelar polimorfismos ligados à resistência. Este tipo de marcador foi denominado DR analogs (disease resistance analogs = análogos de genes de resistência). O domínio conservado NB-ARC mostrado na figura refere-se a uma região ligada a transmissão de sinais que é comum em genes de resistência a doenças em plantas e associados com morte celular programada em animais.

7.5. Piramidização de genes de resistência

Um híbrido de tomateiro moderno deve apresentar a maior quantidade possível (“pacote”) de resistências incorporadas, principalmente contra doenças de solo e de etiologia viral, para as quais os tratamentos curativos são pouco efetivos. Um grande número de fatores de resistência genética presentes em acessos de tomateiro é do tipo monogênico/dominante. Isso possibilita a acumulação de múltiplos fatores

de resistência em um único híbrido. Nesta tarefa, a seleção assistida por marcadores moleculares desempenha um papel fundamental na seleção das linhagens parentais, permitindo “piramidar” genes em linhagens elite e quebrar ligações genéticas com características deletérias e/ou negativa (**Figura 12**).

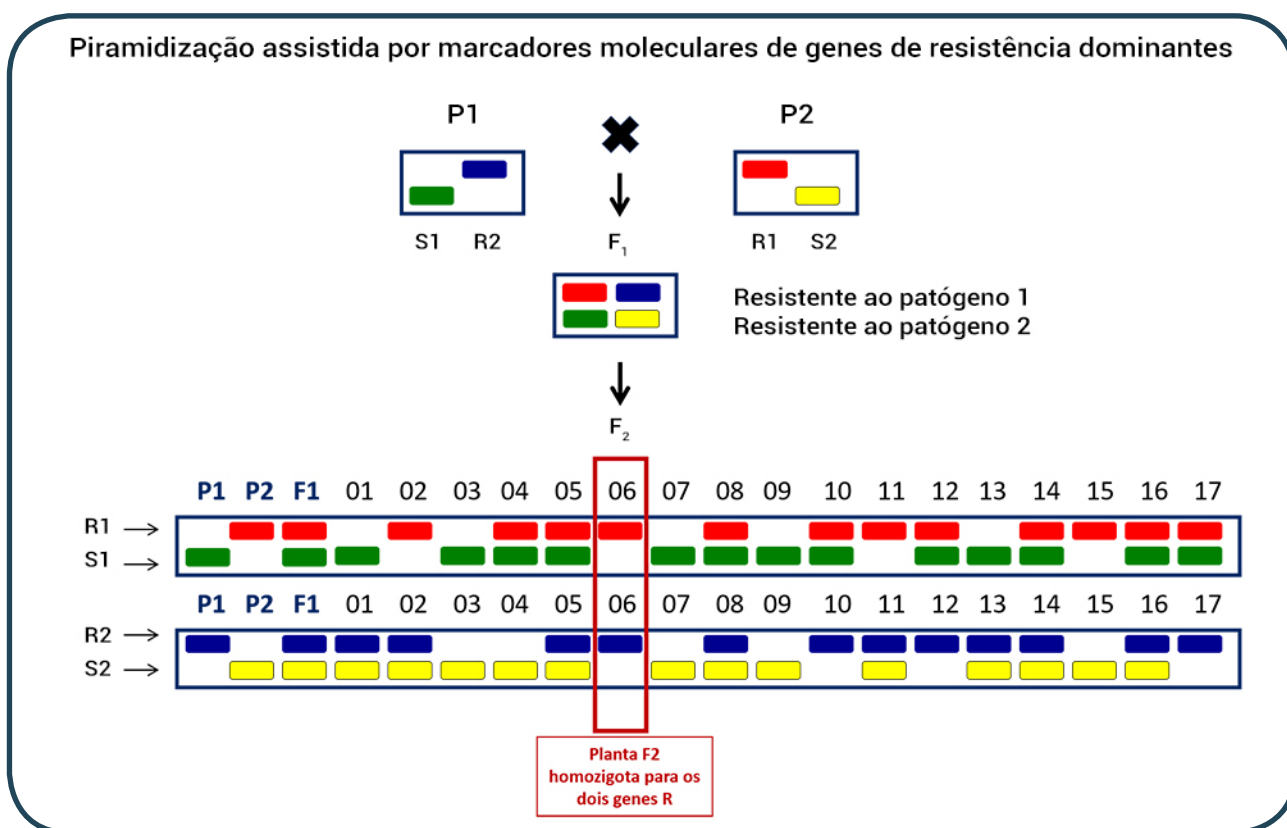


Figura 12. Piramidização de dois genes dominantes de resistência com o auxílio de marcadores moleculares codominantes. Indivíduo 06 (circulado em vermelho) contém os dois marcadores moleculares ligados aos genes de resistência em homozigose (fixados).

7.6. Retrocruzamento acelerado

O retrocruzamento é rotineiramente aplicado em programas de melhoramento para introgressão de genes. Esse método tem por objetivo recuperar o genoma do parental recorrente e, portanto, as características originais do material elite, ao mesmo tempo que incorpora características genéticas com alta herdabilidade, controladas por um ou poucos genes (Fehr, 1987), como por exemplo, fatores de resistência a doenças. Neste contexto, o retrocruzamento acelerado empregando marcadores moleculares é uma estratégia que visa aumentar a eficiência do processo, identificando de maneira precoce indivíduos que apresentam maiores níveis de recuperação do genoma do parental recorrente. A análise de polimorfismos distribuídos ao longo de todo o genoma dos parentais empregados nos cruzamentos envolve a chamada **genotipagem gráfica** onde são selecionados indivíduos com o novo gene de interesse incorporado e que possuem um menor número de segmentos cromossômicos do parental doador (Young & Tanksley, 1989). Este processo é conduzido preferencialmente com marcadores codominantes distribuídos, de preferência, em todos os braços dos cromossomos.

Essa estratégia pode fazer com que os níveis de recuperação do genoma de interesse possam ser rapidamente alcançados equivalendo a “saltos” em relação ao retrocruzamento convencional (**Figura 13**).

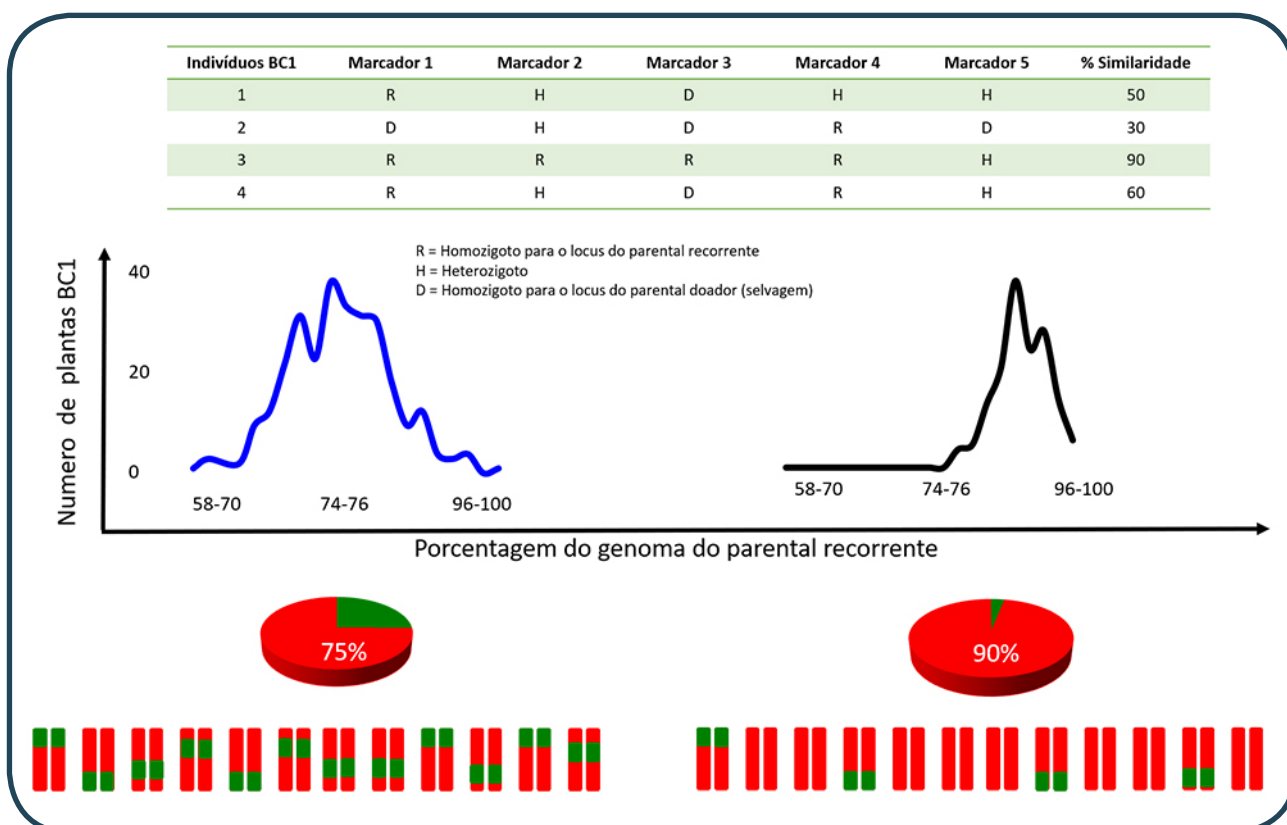


Figura 13. Esquema indicando o “salto” na recuperação do genoma do parental recorrente em sistema de retrocruzamento acelerado assistido por marcadores moleculares. A figura ilustra esse fato com o indivíduo #3 da população segregante resultante do retrocruzamento 1 (back-cross – BC1) que apresentou 90% de similaridade com o genoma do parental recorrente (linha-gem elite) após genotipagem para cinco marcadores distribuídos pelo genoma do tomateiro. Isso indivíduo representa um incremento em comparação com a média teórica de recuperação do genoma do parental recorrente prevista nessa geração (75%).

7.7. Outras aplicações dos marcadores moleculares ligados a fatores de resistência

Marcadores moleculares codominantes ligados aos diferentes genes de resistência estão sendo empregados simultaneamente em sistemas de seleção assistida e para checagem de **pureza genética de híbridos** (Bhavana et al., 2019). A presença de misturas físicas de sementes ou de sementes oriundas apenas da autofecundação do parental feminino de um dado híbrido podem ser prontamente detectadas com esses tipos de marcadores. Os marcadores codominantes também podem ser utilizados para estimar o efeito de **dosagem gênica** na expressão fenotípica da resistência, como demonstrado por Ferro (2013) para o gene *Ty-1* que controla resistência contra begomovírus no tomateiro. Além disso, os marcadores moleculares ligados ou derivados

de genes de resistência são também de grande interesse como potenciais descritores dentro do sistema de **proteção de cultivares** preconizado pela International Union of Protection of Plant Varieties (UPOV). A avaliação de reação das novas potenciais cultivares para doenças nem sempre são simples devido a influência de fatores ambientais, dificultando uma fenotipagem precisa (Arens et al., 2010). Desta forma, dentro do contexto da UPOV os marcadores ligados a fatores de resistência têm o potencial para serem uma alternativa aos testes de patogenicidade, visando cumprir os requisitos necessários para serem incluídos nos testes de DHE (distinguibilidade, homogeneidade e estabilidade) de novas variedades para concessão dos direitos dos melhoristas (Plant Breeders Rights – PBR) (Arens et al., 2010).

8. Exemplos de seleção assistida por marcadores (SAM) moleculares para diferentes patossistemas do tomateiro

A seleção assistida por marcadores consiste na identificação precoce de plantas que apresentam fenótipo desejável pela análise molecular de seu genótipo, o chamado **melhoramento genético acelerado** (Gosal & Wani, 2020) que também está sendo exercitado com o tomateiro (Kaushal et al., 2020). Os três requisitos básicos de um marcador para contribuir para SAM são: **(1)** O marcador deve estar fortemente associado ao gene de interesse; **(2)** A técnica deve ser eficiente para a avaliação de grandes populações e, preferencialmente, baseada em PCR; **(3)** A técnica deve mostrar-se altamente reproduzível entre laboratórios, ser econômica e de fácil condução (Mohan et al., 1997). O desenvolvimento de mapas genéticos extremamente densos tem permitido a localização genômica/isolamento de genes e a identificação de diversos marcadores moleculares com essas propriedades desejadas para emprego em SAM (Barone et al., 2009, Foolad & Panthee, 2012). SAM tem sido utilizada com sucesso para vários caracteres de herança simples (monogênicos), fazendo o processo de seleção mais rápido e eficiente. Vários marcadores moleculares em estreita ligação ou mesmo dentro de alguns desses genes (marcadores funcionais) já estão disponíveis para ser utilizados em seleção assistida. Regiões do genoma que possuem genes controlando a expressão de características complexas (poligênicas = QTLs) também têm sido mapeadas (Tanksley & Fulton, 2007). Os QTLs associados com uma dada característica quantitativa podem estar localizados em diferentes cromossomos e apresentar diferentes contribuições na expressão do fenótipo em estudo. O tomateiro é uma das plantas que serviram de modelo para o isolamento dos primeiros QTLs envolvendo tamanho e massa do fruto (Tanksley & Fulton, 2007). Vários QTLs associados com diferentes caracteres de efeito quantitativo já foram identificados e mapeados tais como: rendimento, qualidade de fruta, tolerância a estresses e **resistência horizontal** a diferentes doenças (Young, 1996). A seguir, vamos apresentar alguns exemplos da utilização de SAM no melhoramento genético do tomateiro.

8.1. Seleção assistida para fatores de resistência contra raças de *Fusarium oxysporum* f. sp. *lycopersici* (FOL)

A doença murcha vascular é causada por raças do fungo *Fusarium oxysporum* f. sp. *lycopersici* (FOL). Quatro genes da série I (*Immunity*, designados como *I-1*, *I-2*, *I-3* e *I-7*) foram caracterizados conferindo tanto resistência do tipo raça-específica quanto resistência de amplo espectro contra isolados de FOL. O primeiro gene identificado (*I-1*) apresenta resistência restrita a isolados de FOL raça 1, tendo sido introgridido da espécie silvestre *Solanum pimpinellifolium* e mapeado no cromossomo 11. Um segundo gene (*I-2*) se mostrou efetivo contra isolados de FOL raça 1, sendo introgridido do acesso *S. pimpinellifolium* PI 126915. O segmento genômico portando o gene *I-2* já foi clonado e isolado (Simons et al., 1998), sendo localizado também no cromossomo 11. Um novo gene de resistência efetivo contra isolados de FOL raça 3 foi introgridido a partir do acesso de *S. pennellii* ‘LA-716’ e se localiza no cromossomo 7 (Panthee & Chen, 2010). Posteriormente foi verificado que esse gene introgridido a partir de um acesso *S. pennellii* (denominado *I-3*) confere resistência a isolados FOL raça 2 e raça 3 (Catanzariti et al., 2015). O gene *I-3* pertence à família de genes do tipo “S-receptor-like kinase” (SRLK), constituindo uma nova classe de genes de resistência (Catanzariti et al., 2015). O gene dominante *I-7* (localizado no cromossomo 8) funciona contra isolados de todas as raças FOL, tendo sido introgridido de *S. pennellii* PI 414773 (Gonzalez-Cendales et al., 2016). O gene *I-7* foi isolado com o auxílio de análise de transcrito (Gonzalez-Cendales et al., 2016) e codifica uma proteína do tipo “leucine-rich repeat receptor-like protein” (LRR-RLP). Marcadores moleculares ainda não estão disponíveis para o gene *I-7*, que foi apenas recentemente clonado (Gonzalez-Cendales et al., 2016). No entanto, sistemas de marcadores moleculares específicos foram desenvolvidos e já se encontram disponíveis para monitorar as regiões genômicas que contêm os fatores de resistência *I-1* (Parmar & Subramanian, 2011), *I-2* (El Mohtar et al., 2007) e *I-3* (Lim et al., 2006; 2008). Um protocolo para detecção simultânea dos genes *I-1*, *I-2* e *I-3* via PCR multiplex já está disponível e pode ser empregado para auxiliar o processo de seleção (Carrer Filho et al., 2016). Um exemplo de detecção do específica do gene *I-3* é ilustrado na **Figura 14**.

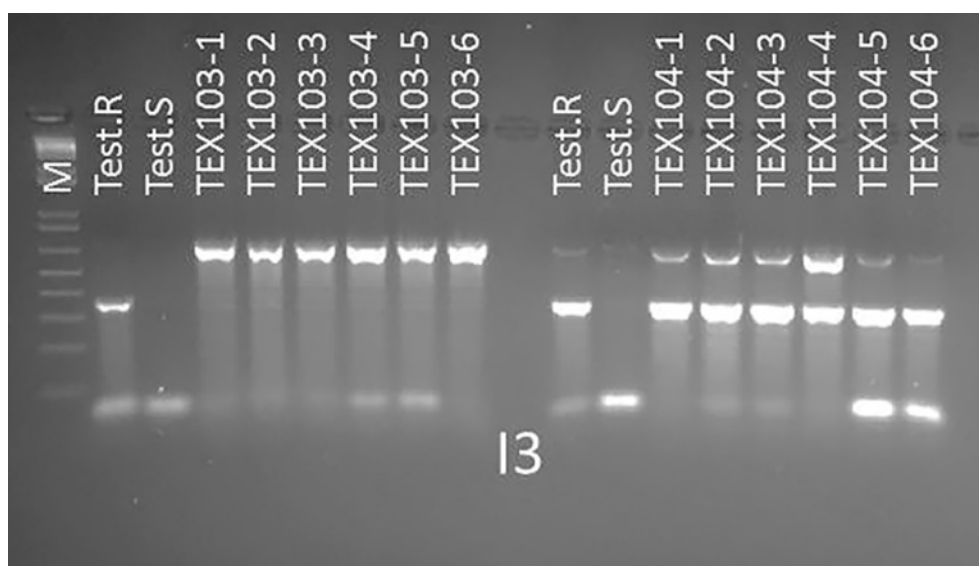


Figura 14. Marcador molecular codominante para detecção do gene *I-3* no cromossomo 7 (introgredido a partir do acesso de *Solanum pennellii* 'LA-716') que confere resistência a isolados *Fusarium oxysporum* f. sp. *lycopersici* classificados dentro das raças 2 e 3. As linhagens TEX104-1 até TEX104-4 são homozigotas resistentes, enquanto as linhagens TEX104-5 e TEX104-6 são heterozigotas para o gene *I-3*. Para os detalhes metodológicos deste marcador ver Barillas et al. (2008).

8.2. Seleção assistida resistência à podridão de co-roa causada por isolados de *Fusarium oxysporum* f. sp. *radicis-lycopersici*

Um gene dominante (*Frl*) de resistência a isolados de *Fusarium oxysporum* f. sp. *radicis-lycopersici* foi introgredido de *S. peruvianum*, estando presente nos acessos PI 126944, PI 128650 e PI 126926. O gene *Frl* está localizado no cromossomo 9 e ligado com os genes/alelos do gene *Tm-2/Tm-2²* que controlam resistência a patótipos de to-mato mosaic virus – ToMV (Fazio et al., 1999). Esse gene ainda não foi clonado e apenas marcadores RAPD e SCAR estão disponíveis para esse locus (Fazio et al., 1999).

8.3. Seleção assistida para fatores de resistência contra a requeima (*Phytophthora infestans*)

O primeiro gene de resistência para *P. infestans* descrito no tomateiro foi o locus *Ph-1* introgredido de um acesso de *S. pimpinellifolium* que se mostrou efetivo contra um

número limitado de raças/variantes do patógeno. Essa resistência foi mapeada no cromossomo 7, mas não se mostra efetiva em condições amplamente favoráveis ao patógeno (Panthee & Chen, 2010). O gene *Ph-2* (também oriundo de *S. pimpinellifolium*) foi descrito posteriormente e apresenta dominância parcial contra alguns isolados de *P. infestans*. Este gene está localizado no cromossomo 10 em estreita associação com os marcadores moleculares TP105 e TG233, mas apresenta certa instabilidade térmica (Panthee & Chen, 2010). O gene de resistência a *P. infestans* mais utilizado em programas de melhoramento é o *Ph-3* (também introgridido de *S. pimpinellifolium*) que apresenta o mais amplo espectro de eficiência. O gene *Ph-3* está localizado no cromossomo 9 em estreita associação com o marcador molecular TG591A (Panthee & Chen, 2010). Mais recentemente, o gene *Ph-3* foi isolado via clonagem posicional (Zhang et al., 2014). A análise de sua sequência apresentou as características típicas de um gene R contendo os domínios conservados CC-NB-LRR. Novos marcadores derivados da própria sequência do gene que estão disponíveis após a clonagem deste fator de resistência (Zhang et al., 2014). Uma análise de um marcador ligado com o gene *Ph-3* é apresentada na **Figura 15**. QTLs controlando uma resistência horizontal de amplo a isolados de *P. infestans* foram também mapeados em cruzamentos envolvendo a fonte original da resistência (*S. habrochaites*) e com linhagens de tomateiro cultivado (Li et al., 2011).

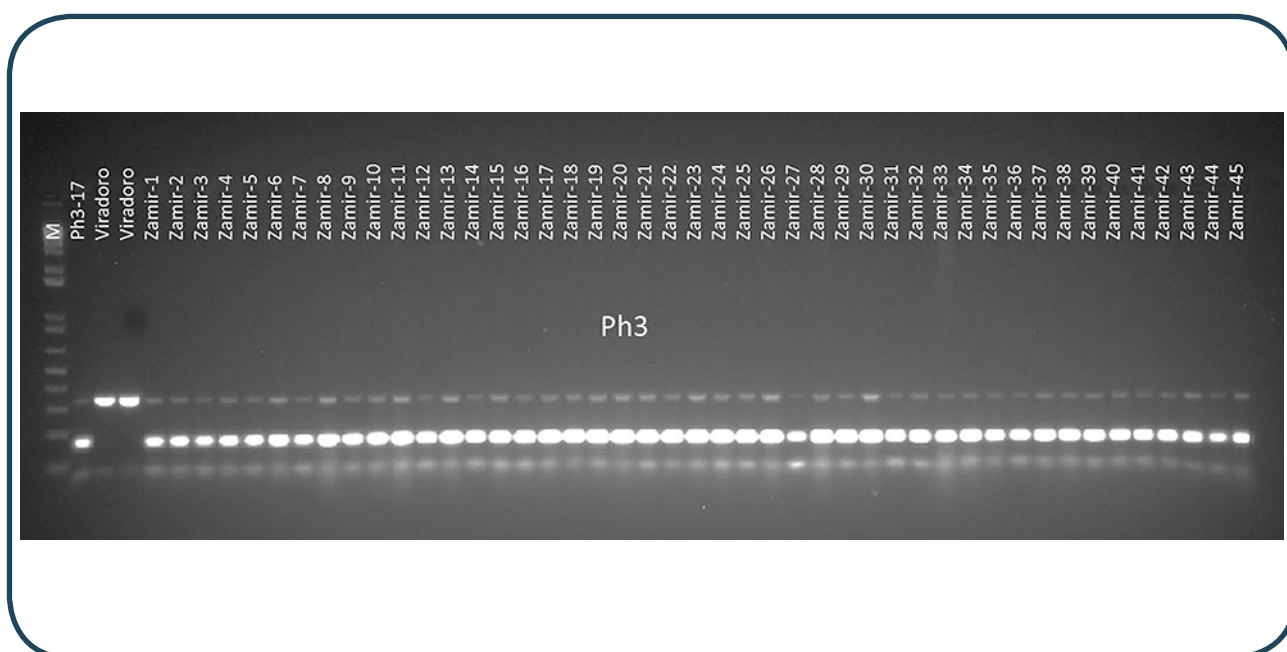


Figura 15. Marcador codominante para detecção do gene *Ph-3* no cromossomo 9 (introgridido de *Solanum pimpinellifolium*), que confere resistência contra um amplo espectro de isolados do oomiceto *Phytophthora infestans* (Zhang et al., 2014). A cultivar 'Viradoro' é suscetível ao oomiceto enquanto que a linhagem *Ph3-17* e o híbrido BRS Zamir são resistentes. Esse marcador codominante também serve para checar a pureza do híbrido BRS Zamir. No exemplo, dez plantas desse híbrido foram amostradas e 100% delas se mostraram heterozigotas para o gene *Ph-3*.

8.4. Seleção assistida para fatores de resistência à mancha foliar de *Cladosporium*

Alelos e genes (*Cf*-1 até *Cf*-9) que controlam resistência a diferentes raças de *Pas-salora fulva* (= *Cladosporium fulvum*). Os genes *Cf*-2 e *Cf*-5 estão no cromossomo 6 ligados em associação com o gene *Mi*-1.2 de resistência aos nematoides das galhas (Dixon et al., 1996). Por sua vez, os genes *Cf*-1, *Cf*-4 e *Cf*-9 estão localizados no cromossomo 1 (Jones et al., 1994; Thomas et al., 1997). Todos esses genes já foram clonados bem como os genes de avirulência correspondentes no patógeno (Medina et al., 2015). Uma análise de um marcador associado com o gene *Cf*-9 é apresentada na **Figura 16** junto com um gel com o marcador molecular ligado gene *Mi*-1.2 (que funciona também como marcador para a presença dos genes *Cf*-2 e *Cf*-5) devido ao fato desse bloco gênico residir em um segmento cromossômico com uma frequência de recombinação extremamente reduzida (Rossi et al., 1998; Vos et al., 1998).

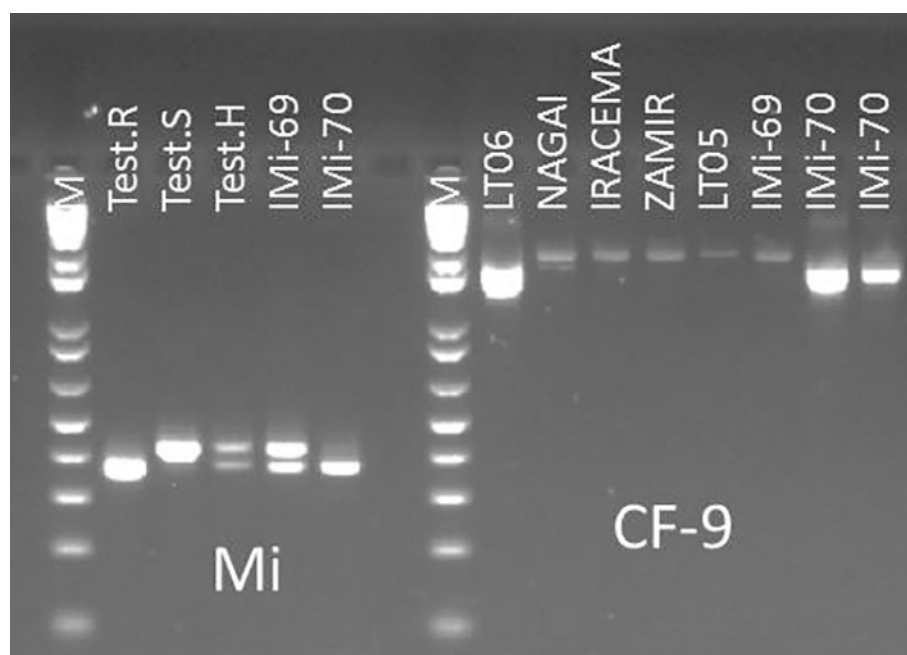


Figura 16. Uma análise de um marcador dominante associado com o gene *Cf*-9 (localizado no cromossomo 1) apresentada junto de um gel com o marcador molecular codominante ligado gene *Mi*-1.2. Os genes/alelos *Cf*-2 e *Cf*-5 estão no cromossomo 6 ligados em associação com o gene *Mi*-1.2 de resistência aos nematoides das galhas. Marcador ligado ao gene *Mi*-1.2 é detectado diretamente via PCR devido a presença de um INDEL entre os sítios de anelamento dos primers Mi23F: 5'-TGG-AAA-AAT-GTT-GAA-TTT-CTT-TTG-3' e Mi23R: 5'-GCA-TAC-TAT-AT-G-GCT-TGT-TTA-CCC-3' (ver detalhes metodológicos em Bhavana et al., 2019). A linhagem LT-06 apresenta o gene *Cf*-9 assim como a linhagem IMI-70. As demais linhagens não possuem o gene. A linhagem IMI-70 também é homozigota para o gene de resistência *Mi*-1.2 enquanto a linhagem IMI-69 é heterozigota.

8.5. Seleção assistida para fatores de resistência contra *Verticillium dahliae* raça 1

Resistência a isolados de *Verticillium* foi inicialmente identificada em 1932 no acesso *S. lycopersicum* var. *cerasifore* ‘Peru Wild (Kawchuk et al., 2001). Esta resistência é raça 1-específica, apresentando herança simples, monogênica e dominante (Kawchuk et al., 2001). A resistência derivada de ‘Peru Wild’ é condicionada por um locus complexo composto por dois genes *Ve-1* e *Ve-2* que se localizam no cromossomo 9 do tomateiro (Diwan et al., 1999; Kawchuk et al., 2001). Experimentos posteriores a clonagem desse locus, indicaram que apenas o gene *Ve-1* é o responsável pela expressão da resistência (Fradin et al., 2009). Um marcador codominante ligado ao gene *Ve-1* é apresentado na **Figura 17**. Esse sistema de marcadores não apresenta estabilidade e pode ser resultado de depósitos de sequência errôneos como sugerido por Arens et al. (2010).

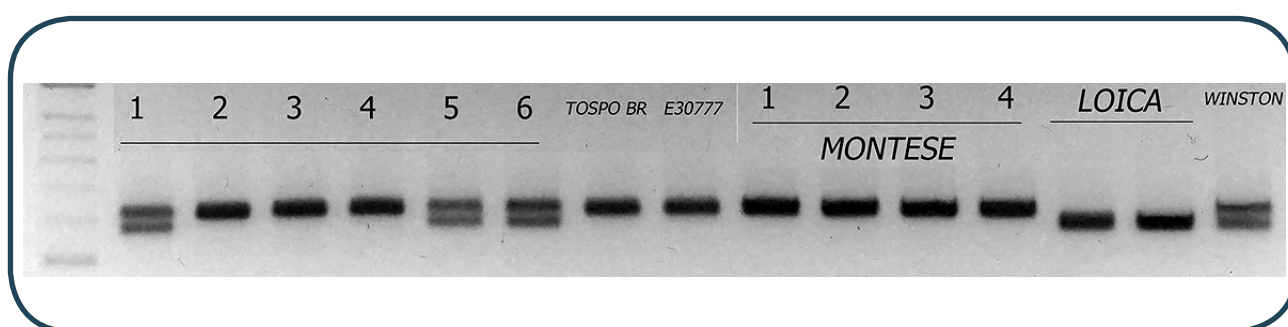
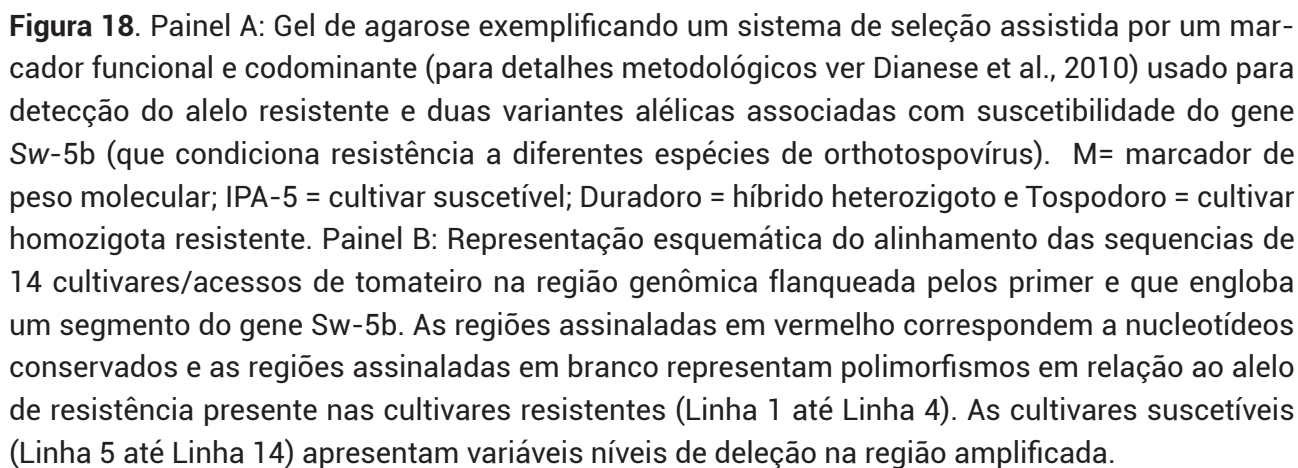


Figura 17. Gel de agarose (2%) para detecção do marcador molecular codominante derivado de informação genética do locus complexo (*Ve-1* e *Ve-2*) no cromossomo 9 que controla resistência a isolados de *Verticillium dahliae* raça 1 no tomateiro (Kawchuk et al., 2001; Arens et al., 2010). Dados recentes indicam que apenas o gene *Ve-1* é o responsável pela expressão da resistência ao fungo. O híbrido BRS Montese é resistente a *V. dahliae* raça 1 (banda alta), enquanto a cultivar ‘Loica’ é suscetível (banda baixa). O híbrido ‘Winston’ é heterozigoto para o gene *Ve-1* (presença das duas bandas).

8.6. Seleção assistida para resistência a orthotospovirus

As três espécies predominantes de *Orthotospovirus* que causam a doença conhecida como “vira-cabeça” do tomateiro no Brasil são: tomato spotted wilt virus (TSWV), tomato chlorotic spot virus (TCSV) e groundnut ringspot virus (GRSV). O gene *Sw-5b* confere resistência de tipo imunidade para quase todas as variantes genéticas dentro do gênero e tem sido o fator mais utilizado no desenvolvimento de cultivares resistentes (Oliveira et al., 2018). Esse locus foi introgridido de *S. peruvianum* e apresenta um amplo espectro de ação contra distintas espécies de *Tospovirus* que infectam tomateiro (Dianese et al., 2011). O gene *Sw-5b* é um típico membro da família de genes R do classe CC-NBS-LRR (Brommonschenkel et al., 2000). O gene *Sw-5b* foi mapeado no



8.7. Seleção assistida para resistência a geminivírus (=begomovírus)

Os programas de melhoramento dispõem de distintas fontes de resistência contra begomovirose, muitas delas mapeadas e com os seus produtos gênicos caracterizados (**Tabela 1**). O gene *Ty-1* controla resistência parcial ao begomovírus *Tomato yellow leaf curl virus* (De Castro et al., 2007). O gene *Ty-1* também é efetivo contra espécies virais de genoma bipartido do Brasil (Boiteux et al., 2012b), sendo localizado no cromossomo 6 em repulsão com o gene *Mi-1.2* (De Castro et al., 2007). Análise com um sistema de marcador molecular ligado ao locus *Ty-1* está ilustrado na **Figura 19**. Trabalhos de mapeamento genético conduzidos por Verlaan et al. (2011) permitiram isolar o gene *Ty-1* e confirmar que ele codifica uma RNA polimerase dependente de RNA (RDR) pertencente ao tipo RDRy, representando uma nova classe de fatores de resistência inteiramente nova que confere esse fenótipo através da intensificação dos níveis do silenciamento gênico transcricional (Butterbach et al., 2014). O gene *Ty-2* (localizado no cromossomo 11) foi introgridido de *S. habrochaites* (Yang et al., 2014) e funciona também contra espécies de genoma bipartido do Brasil (Boiteux et al., 2012b). Recentemente se determinou que o gene *Ty-2* pertence à classe de genes R contendo os domínios conservados NB-LRR (Yamaguchi et al., 2018). O gene *Ty-3* foi introgridido dos acessos de *S. chilense* 'LA-1932' e 'LA-2779' (Hutton et al., 2012) e está ligado ao locus *Ty-1* no cromossomo 6 (Ji et al., 2007). Estudos moleculares demonstraram que os genes *Ty-1* e *Ty-3* são muito provavelmente alélicos (Verlaan et al., 2013). O gene *Ty-4* foi introgridido de *S. chilense* 'LA-1932', localizado no cromossomo 3 (Ji et al., 2009). Estudos de mapeamento e validação funcional do gene recessivo *ty-5* (Hutton et al., 2012) demonstraram que este fator corresponde a um gene *Pelota* (envolvido na fase de reciclagem do ribossomo na síntese de proteína). Silenciamento desse gene em uma linhagem suscetível tornou as plantas transgênicas altamente resistentes ao TYLCV. O gene *Pelo* oferece, portanto, uma rota alternativa para promover a resistência ao TYLCV e outros vírus relacionados ou não (Lapidot et al., 2015; Wang et al., 2018). O gene recessivo *tcm-1* condiciona um amplo espectro de resistência contra distintas espécies de begomovírus com genoma bipartido e monopartido (Giordano et al., 2005) e está localizado no cromossomo 6 (Machado, 2013).

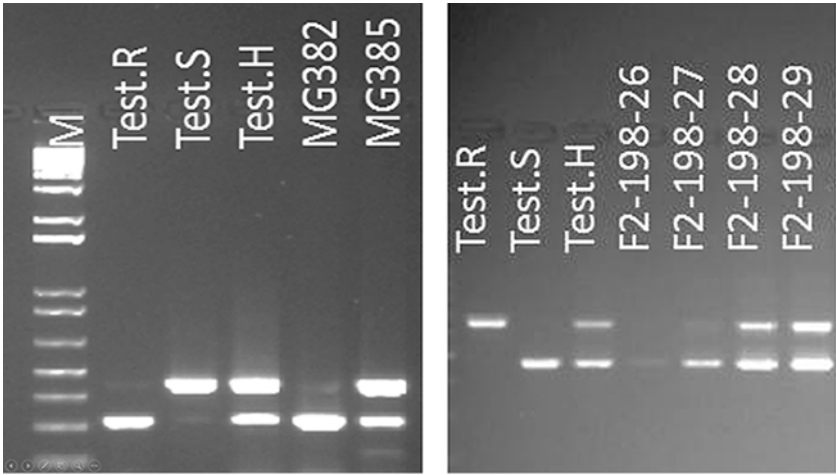


Figura 19. Análise com marcadores codominantes ligados ao locus *Ty-1* (marcador do tipo CAPS) e ao locus *Ty-3* (marcador codominante obtido diretamente via PCR devido a presença de um INDEL na região interna ao sítio de anelamento dos primers). O gene *Ty-1* controla resistência parcial ao begomovírus *Tomato yellow leaf curl virus* e contra espécies virais de genoma bipartido que ocorrem no Brasil. O gene *Ty-3* foi introgridido dos acessos de *S. chilense* 'LA-1932' e 'LA-2779' está ligado ao locus *Ty-1* no cromossomo 6. Estudos moleculares demonstraram que os genes *Ty-1* e *Ty-3* são muito provavelmente alélicos. Para detalhes metodológicos ver Machado (2013).

Tabela 1. Genes de resistência a espécies de *Begomovirus* utilizados em sistemas de seleção assistida por marcadores no melhoramento genético do tomateiro.

Gene	Localização cromossômica	Mecanismo de resistência	Espectro de eficiência (contra quais begomovírus a resistência foi avaliada)	Fontes de resistência
<i>Ty-1</i>	6 (Verlaan et al., 2013)	DFDGD-class RNA-dependent RNA polymerase (Verlaan et al., 2013)	TYLCV (Vidasky et al., 2008); ToCMoV , ToRMV , ToSRV ToYVSV (Boiteux et al., 2012b).	<i>S. chilense</i> LA 1969
<i>Ty-2</i>	11 (Yang et al., 2014)	Proteína contendo domínios <i>nucleotide-binding site</i> (NB) & <i>leucine rich repeat</i> (LRR) (Yamaguchi et al., 2018)	TYLCV (Yamaguchi et al., 2018); ToSRV (Boiteux et al., 2012b)	<i>S. habrochaites</i> B6013 e H-24
<i>Ty-3</i>	6 (Verlaan et al., 2013)	Variante alélica de <i>Ty-1</i> (Verlaan et al., 2013)	TYLCV & ToMoV (Vidasky et al., 2008)	<i>S. chilense</i> LA 1932
<i>Ty-4</i>	3 (Ji et al., 2009)	Ainda não caracterizada	TYLCV & ToMoV (Vidasky et al., 2008)	<i>S. chilense</i> LA 1932
<i>ty-5</i>	4 (Lapidot et al., 2015)	Gene codificador de um homólogo do mRNA surveillance factor <i>Pelota</i> (<i>Pelo</i>) (Lapidot et al., 2015).	TYLCV (Gill et al., 2019). Não funciona para isolados de ToMoV da Flórida (Gill et al., 2019).	<i>S. peruvianum</i>

Gene	Localização cromossômica	Mecanismo de resistência	Espectro de eficiência (contra quais begomovírus a resistência foi avaliada)	Fontes de resistência
Ty-6	10 (Gill et al., 2019)	Ainda não caracterizada	TYLCV & ToMoV (Gill et al., 2019)	<i>S. chilense</i>
tcm-1	6 (Machado, 2013)	Menor acumulação & replicação viral (Boiteux et al., 2012b)	ToCMoV (Giordano et al., 2005); ToSRV; ToRMV; ToYVSV (Machado, 2013); ToYLCV (Garcia-Cano et al., 2008)	'TX-468'
tgr-1	6 (Bian et al., 2007)	Ainda não caracterizada	TYLCV (Bian et al., 2007)	'FLA 653'

8.8. Seleção assistida para resistência à mancha-bacteriana

Um complexo de espécies e patovares dentro do gênero *Xanthomonas* está associado com a mancha-bacteriana no tomateiro. As diferentes espécies/patovares de *Xanthomonas* são classificadas em cinco raças (T1 até T5) de acordo a sua capacidade (ao não) de induzir reações de hipersensibilidade em diferentes genótipos de referência. A correspondência raça/espécie é a seguinte: *X. vesicatoria* (raça T2), *X. euvisicatoria* pv. *euvisicatoria* (raça T1), *X. euvisicatoria* pv. *perforans* (raças T3, T4 e T5) e *X. cynarae* pv. *gardneri* (também raça T2). Estas bactérias podem atacar todas as partes vegetativas da planta, reduzindo a produtividade pela destruição da folhagem e queda de flores e frutos em formação. Os danos diretos e indiretos provocados por esse grupo de patógenos também têm grandes efeitos na redução da qualidade do produto processado. Diferentes fatores de resistência e QTLs com efeitos importantes contra diferentes espécies/raças já foram descritos nesse patossistema. Alguns desses fatores apresentam marcadores moleculares ligados o que possibilita o monitoramento da região de interesse, facilitando o processo de “piramidização” (Yang & Francis, 2005). Três QTLs (*Rx-1*, *Rx-2* e *Rx-3*) controlam tolerância/hipersensibilidade (HS) a isolados de *X. euvisicatoria* pv. *euvisicatoria* raça T1 no acesso ‘Hawaii-7998’. Os QTLs *Rx-1* e *Rx-2* estão localizados no cromossomo 1 (Hutton, 2008). QTL *Rx-3* está em repulsão com o “cluster” *Pto/Fen/Prf* no cromossomo 5 (Yang & Francis, 2005). Um gene dominante e raça-T4 específico (denominado *Bs-4*) introgridido de *S. pennellii* ‘LA-716’ foi localizado no cromossomo 5. O gene *Bs-4* confere resistência a isolados de *X. perforans* raça T4 (Ballvora et al., 2001). Análise usando todos os marcadores polimórficos associados com o gene *Xv3* (que controla uma resposta hipersensível a isolados de *X. euvisicatoria* pv. *perforans* raça T3) indicaram que a resistência está localizada no cromossomo 11. Os testes de alelismo foram realizados em cruzamentos entre linhagens portadoras do gene *Xv3* (derivado da linhagem ‘Hawaii 7981’), do gene *Rx-4* (derivado do PI 128216) e da resistência derivada do acesso PI 126932. Esses testes indicaram que os loci que condicionam HR para *X. euvisicatoria* pv. *perforans* raça T3 estão ligados dentro de um intervalo de 0,1 cM ou são alélicos ou são o mesmo gene (Wang et al., 2011).

8.9. Seleção assistida para resistência à murcha bacteriana (*Ralstonia solanacearum* e *R. pseudosolanacearum*)

As linhagens ‘Hawaii 7997’ e ‘Hawaii 7996’ apresentam os melhores níveis de resistência e os mais amplos espectros de ação contra diferentes isolados de *R. solanacearum* e *R. pseudosolanacearum* (Carmeille et al., 2006; Geethanjali et al., 2010). Quatro QTLs (denominados de *Bwr*) foram identificados no acesso ‘Hawaii 7996’ e numerados de acordo com a sua posição cromossômica. Essas regiões explicaram entre 3,2% e 29,8% da variação fenotípica da resistência. Um QTL de maior efeito está no cromossomo 6 (*Bwr*-6) e um de menor efeito está no cromossomo 3 (*Bwr*-3) (Carmeille et al., 2006; Geethanjali et al., 2010). Os QTLs *Bwr*-4 e *Bwr*-8 foram detectados apenas sob elevada temperatura (Carmeille et al., 2006).

8.10. Seleção assistida para resistência aos nematoides-das-galhas (gênero *Meloidogyne*)

O tomateiro é uma das mais importantes hospedeiras dos nematoides das galhas. *Meloidogyne incognita*, *M. javanica* e *M. arenaria* são as espécies de mais ampla distribuição geográfica em cultivos de tomateiro no mundo. No Brasil, levantamentos conduzidos indicaram a presença das seguintes espécies infectando o tomateiro: *M. javanica*, *M. incognita* (EST I1 e I2), *M. ethiopica*, *M. enterolobii* (= *M. mayaguensis*) e *M. morocciensis*. A resistência genética se constitui na estratégia mais eficiente e de maior sustentabilidade para o controle dos nematoides das galhas no tomateiro. Atualmente a resistência disponível para o melhoramento genético do tomateiro é baseada em um gene dominante *Mi*-1.2 que foi introgridido do acesso da espécie selvagem *S. peruvianum* PI 128657. Esse gene se localiza no braço curto do cromossomo 6 e tem sido amplamente empregado em cultivares e híbridos comerciais uma vez que confere resistência a distintas populações/raças de *M. incognita*, *M. javanica* e *M. arenaria* e mais dez outras espécies de *Meloidogyne* (Gabriel et al., 2020). SAM para o gene *Mi*-1.2 foi inicialmente conduzida com sucesso utilizando marcadores bioquímicos (isoenzimas) codificadas pelos diferentes alelos do gene da enzima fosfatase ácida 1 (**Figura 6**). Posteriormente, o gene *Mi*-1.2 foi isolado via clonagem posicional (Rossi et al., 1998; Vos et al., 1998). Análises genômicas adicionais indicaram que o gene *Mi*-1.2 faz parte de um “cluster” de sete cópias homólogas que estão presentes nos genomas de acessos suscetíveis e resistentes. Todos os membros do “cluster” do gene *Mi*-1.2 são classificados na família de genes R contendo os elementos conservados CC–NB–LRR. No entanto, a função destes outros genes homólogos do “cluster” permanece ainda por ser elucidada. Marcadores moleculares codominantes com estreita ligação ao gene *Mi*-1.2 foram desenvolvidos e estão sendo empregados em sistemas aplicados de seleção assistida e para checagem de pureza genética de híbridos (Bhavana et al., 2019). Uma análise de um desses marcadores codominantes está ilustrada na **Figura 16**. O gene *Mi*-1.2 apresenta um interessante efeito pleiotrópico, controlando resistência também contra populações do afídeo *Macrosiphum euphorbiae* e da mosca-branca *Bemisia tabaci*. Como men-

cionado, outro aspecto positivo associado ao locus contendo o gene *Mi-1.2* é a presença de uma estreita ligação como os genes *Cf-2* e *Cf-5* (que controlam resistência contra algumas raças do fungo *Cladosporium*). Entretanto, o gene *Mi-1.2* não é efetivo em temperaturas acima de 28° C e também não tem efeito contra populações de *M. hapla* e *M. enterolobii*. Além disso, já existem registro de isolados virulentos de *M. incognita* e *M. javanica*. Outro problema em termos de melhoramento genético é o fato do gene *Mi-1.2* se encontrar em uma região genômica em forte repulsão com o gene dominante *Ty-1* e o gene recessivo *tcm-1* (que controlam resistência a diferentes espécies de *Begomovirus*), dificultando, desta forma, a obtenção de linhagens puras contendo esses fatores em homozigose. Mais recentemente, um novo gene de resistência denominado de *Mi-9* (introgridido da espécie selvagem *S. arcanum* ‘LA 2157’) foi clonado e caracterizado. O gene *Mi-9* é um homólogo do gene *Mi-1.2* e está também localizado no cromossomo 6. Um aspecto diferencial do *Mi-9* é o fato desse gene apresentar maior estabilidade térmica que o gene *Mi-1.2* (Jablonska et al., 2007).

9. Considerações finais

Como mostrado na **Figura 5** e exemplificado ao longo desse capítulo, os marcadores moleculares representam uma tecnologia crucial no melhoramento genético de nova geração e na biotecnologia do tomateiro. Embora o tomateiro tenha sido utilizado como exemplo, todos os aspectos positivos dos marcadores moleculares descritos aqui já estão sendo capitalizados pelos programas de melhoramento genético das principais hortaliças e de todas as importantes culturas alimentares, viabilizando sistemas de SAM e o “melhoramento acelerado” de novas cultivares e híbridos (Gosal & Wani, 2020).

10. Referências bibliográfica

- AKKAYA, M.S.; BHAGWAT, A.A.; CREGAN, P.B. Length polymorphisms of simple sequence repeat DNA in soybean. **Genetics**, v.132, p.1131–1139, 1992.
- ALFENAS, A.C. **Eletroforese de isoenzimas e proteínas afins: Fundamentos e aplicações em plantas e microorganismos**. Viçosa: Editora UFV. 1998. 574 p.
- ARENS, P. et al. Development and evaluation of robust molecular markers linked to disease resistance in tomato for distinctness, uniformity and stability testing. **Theoretical and Applied Genetics**, v. 120, p. 655–664, 2010.
- BAGGS, E.; DAGDAS, G.; KRASILEVA, K.V. NLR diversity, helpers and integrated domains: Making sense of the NLR Identity. **Current Opinion in Plant Biology**, v. 38, p. 59–67, 2017.
- BALLVORA, A.; SCHORNACK, S.; BAKER, B.J.; GANAL, M.; BONAS, U.; LAHAYE, T. Chromosome landing at the tomato *Bs4* locus. **Molecular Genetics and Genomics**, v. 266, 639–645, 2001.

- BARILLAS, A.C.; MEJIA, L.; SANCHEZ-PEREZ, A.; MAXWELL, D.P. CAPS and SCAR markers for detection of *I*-3 gene introgression for resistance to *Fusarium oxysporium* f. sp. *lycopersici* race 3. **Report of the Tomato Genetics Cooperative**, v. 58, p.11–17, 2008.
- BARONE, A.; DI MATTEO, A.; CARPUTO, D.; FRUSCIANTE, L. High-throughput genomics enhances tomato breeding efficiency. **Current Genomics**, v. 10, p. 1–9, 2009.
- BERED, F.; NETO, J.F.B.; CARVALHO, F.I.F. Marcadores moleculares e sua aplicação no melhoramento genético de plantas. **Ciência Rural**, v. 27, 513–520, 1997.
- BHAVANA, P. et al. Identification of resistance in tomato against root knot nematode (*Meloidogyne incognita*) and comparison of molecular markers for *Mi* gene. **Australasian Plant Pathology**, v. 48, p. 93–100, 2019.
- BOITEUX, L.S.; FONSECA, M.E.N.; VIEIRA, J.V.; PEREIRA-CARVALHO, R.C. Melhoramento para resistência a doenças virais. In: BORÉM A; FRITSCH-NETO, R. (eds). **Melhoramento de Plantas para Condições de Estresses Bióticos**. Visconde de Rio Branco, MG: Editora Suprema, 2012b. p. 89-119.
- BOITEUX, L.S.; FONSECA, M.E.N.; GONZÁLEZ-ARCOS, M. Melhoramento de tomate para processamento industrial. In: NICK, C.; BORÉM, A. (eds). **Melhoramento de Hortaliças**. Viçosa, MG: Editora UFV, 2016. p. 432–464.
- BOMBARELY, A. et al. The Sol genomics network (solgenomics.net): Growing tomatoes using Perl. **Nucleic Acids Research**, v. 39, p. D1149–D1155, 2010.
- BOTSTEIN, D.; WHITE, R.L.; SKOLNICK, M.; DAVIS, R.V. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, v. 32, p. 314–331, 1980.
- BROMMONSCHENKEL, S.H.; FRARY, A.; TANKSLEY, S.D. The broad-spectrum tospovirus resistance gene *Sw-5* of tomato is a homolog of the root-knot nematode resistance gene *Mi*. **Molecular Plant-Microbe Interactions**, v.13, p. 1130–1138, 2000.
- BUTTERBACH, P. et al.. *Tomato yellow leaf curl virus* resistance by *Ty*-1 involves increased cytosine methylation of viral genomes and is compromised by cucumber mosaic virus infection. **Proceedings of the National Academy of Sciences USA**, v. 111, p. 12942–12947, 2014.
- CAETANO-ANOLLÉS, G.; BASSAM, B.J.; GRESSHOFF, P.M. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. **Biotechnology**, v. 9, p. 553–556, 1991.
- CARMEILLE, A. ET AL. Identification of QTLs for *Ralstonia solanacearum* race 3-phylo-type II resistance in tomato. **Theoretical and Applied Genetics**, v. 113, p. 110–121, 2006.
- CARRER FILHO, R. et al. Detecção simultânea de fatores de resistência à murcha de fusário do tomateiro por meio de PCR multiplex. **Pesquisa Agropecuária Brasileira**, v. 51, p. 925–932, 2016.
- CATANZARITI, A.M.; LIM, G.T.; JONES, D.A. The tomato $\overline{I}3$ gene: A novel gene for resistance to *Fusarium* wilt disease. **New Phytologist**, v. 207, p. 106–118, 2015.
- CHEN, A.L. et al. Reassessment of QTLs for late blight resistance in the tomato accession L3708 using a restriction site associated DNA (RAD) linkage map and highly aggressive isolates of *Phytophthora infestans*. **PloS One**, v. 9, p. e96417, 2014.
- COLLARD, B.C.Y.; JAHUFER, M.Z.Z.; BROUWER, J.B.; PANG, E.C.K. An introduction to markers, quantitative loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. **Euphytica**, v.142, p.169–196, 2005.
- DE CASTRO, A.P.; BLANCA, J.M.; DÍEZ, M.J.; VINALS, F.N. Identification of a CAPS marker tightly linked to the *Tomato yellow leaf curl* disease resistance gene *Ty*-1 in tomato. **European Journal of Plant Pathology**, v. 117, p. 347–356, 2007.

- DIANESE, E.C. et al. Development of a locus-specific, co-dominant SCAR marker for assisted-selection of the *Sw-5* (*Tospovirus* resistance) gene cluster in a wide range of tomato accessions. **Molecular Breeding**, v. 25, p. 133–142, 2010.
- DIANESE, E. C.; FONSECA, M.E.N.; INOUE-NAGATA, A. K.; RESENDE, R. O.; BOITEUX, L.S. Search in *Solanum* (section *Lycopersicon*) germplasm for sources of broad-spectrum resistance to four *Tospovirus* species. **Euphytica**, v. 180, p. 307–319, 2011.
- DIWAN, N.; FLUHR, R.; ESHED, Y.; ZAMIR, D.; TANKSLEY, S.D. Mapping of *Ve* in tomato: A gene conferring resistance to the broad-spectrum pathogen, *Verticillium dahliae* race 1. **Theoretical and Applied Genetics**, v. 98, p. 315–319, 1999.
- DIXON, M.S.; JONES, D.A.; KEDDIE, J.S.; THOMAS, C.M.; HARRISON, K. The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. **Cell**, v. 84, p. 451–459, 1996.
- EL MOHTAR, C.A.; ATAMIAN, H.S.; DAGHER, R.B.; ABOU-JAWDAH, Y.; SALUS, M.S.; MAXWELL, D.P. Marker-assisted selection of tomato genotypes with the *I-2* gene for resistance to *Fusarium oxysporum* f. sp. *lycopersici* race 2. **Plant Disease**, v.91, p.758–762, 2007.
- EMMANUEL, E.; LEVY, A.A. Tomato mutants as tools for functional genomics. **Current Opinion in Plant Biology**, v. 5, p. 112–117, 2002.
- ELSHIRE, R.J. et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. **PloS One**, v. 6, p. e19379, 2011.
- ESHED, Y.; ZAMIR, D. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. **Genetics**, v. 141, p.1147–1162, 1995.
- FAZIO, G; STEVENS, M.R.; SCOTT, J.W. Identification of RAPD markers linked to *Fusarium* crown and root rot resistance (*Fr1*) in tomato. **Euphytica**, v.105, p. 205–210, 1999.
- FEHR, W.F. **Principles of cultivar development: Theory and technique**. Macmillan, New York. 1987
- FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3ed. Brasília: Embrapa-Cenargen, 1998. 220 p.
- FERRO, D. D. X. **Estudos com o locus *Ty-1* do tomateiro e busca por novos marcadores moleculares para tolerância ao *Tomato severe rugose virus***. Dissertação de Mestrado. Universidade de Brasília (UnB), 2013.
- FOOLAD, M.R.; PANTHEE, D.R. Marker-assisted selection in tomato breeding. **Critical Reviews in Plant Sciences**, v. 31, p. 93–123, 2012.
- FRADIN, E.F. et al. Genetic dissection of *Verticillium* wilt resistance mediated by tomato *Ve1*. **Plant Physiology**, v. 150, p. 320–332, 2009.
- GABRIEL, M.; KULCZYNSKI, S.M.; MUNIZ, M.F.; BOITEUX, L.S.; CARNEIRO, R.M. Reaction of a heterozygous tomato hybrid bearing the *Mi1.2* gene to 15 *Meloidogyne* species. **Plant Pathology**, v. 69, p. 944–952, 2020.
- GEETHANJALI, S.; CHEN, K.Y.; PASTRANA, D.V.; WANG, J.F. Development and characterization of tomato SSR markers from genomic sequences of anchored BAC clones on chromosome 6. **Euphytica**, v. 173, p. 85–97, 2010
- GIORDANO, L.B.; SILVA-LOBO, V.L.; SANTANA, F.M.; FONSECA, M.E.N.; BOITEUX, L.S. Inheritance of resistance to the bipartite Tomato chlorotic mottle begomovirus derived from *Lycopersicon esculentum* cv. ‘Tyking’. **Euphytica**, v. 143, p. 27–33, 2005.

- GONZÁLEZ-ARCOS, M. et al. A loss-of-function allele of a TAC1-like gene (SlTAC1) located on tomato chromosome 10 is a candidate for the *Erectoid leaf* (*Erl*) mutation. **Euphytica**, v. 215, p. 95, 2019.
- GONZALEZ-CENDALES, Y.; CATANZARITI, A.M.; BAKER, B.; McGRATH, D.J.; JONES, D.A. Identification of *I-7* expands the repertoire of genes for resistance to Fusarium wilt in tomato to three resistance gene classes. **Molecular Plant Pathology**, v. 17, p. 448–463, 2016.
- GOSAL, S.S.; WANI, S.H. Accelerated Plant Breeding. v. 2 (Vegetable Crops). Switzerland: Springer Nature Cham, 2020. 457 p.
- GUPTA, P. et al. Reanalysis of genome sequences of tomato accessions and its wild relatives: Development of tomato genomic variation (TGV) database integrating SNPs and INDELs polymorphisms. **bioRxiv** (no prelo), 2020.
- HAMILTON, J.P. et al. Single nucleotide polymorphism discovery in cultivated tomato via sequencing by synthesis. **The Plant Genome**, v. 5, p. 17–29, 2012.
- HELENTJARIS, T.; SLOCUM, M.; WRIGHT, S.; SCHAEFER, A.; NIENHUIS J. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. **Theoretical and Applied Genetics**, v. 72, p. 761–769, 1986.
- HUTTON, S.F. **Inheritance and mapping of resistance to bacterial spot race T4 (*Xanthomonas perforans*) in tomato, and its relationship to race T3 hypersensitivity, and inheritance of race T3 hypersensitivity from PI 126932**. Ph.D. Thesis, University of Florida, 2008.
- HUTTON, S.F.; SCOTT, J.W.; SCHUSTER, D.J. Recessive resistance to *Tomato yellow leaf curl virus* from the tomato cultivar Tyking is located in the same region as *Ty-5* on chromosome 4. **Hort-Science**, v. 47, p. 324–327, 2012.
- JABLONSKA, B. ET AL. The *Mi-9* gene from *Solanum arcanum* conferring heat-stable resistance to root-knot nematodes is a homolog of *Mi-1*. **Plant Physiology**, v. 143, p. 1044–1054, 2007.
- JI, Y.; SCHUSTER, D.J.; SCOTT, J.W. *Ty-3*, a begomovirus resistance locus near the *Tomato yellow leaf curl virus* resistance locus *Ty-1* on chromosome 6 of tomato. **Molecular Breeding**, v. 20, p. 271–284, 2007.
- JI Y, SCOTT J, SCHUSTER D. Molecular mapping of *Ty-4*, a new Tomato yellow leaf curl virus resistance locus on chromosome 3 of tomato. **Journal of the American Society for Horticultural Science**, v. 134, p. 281–288, 2009.
- JONES, D.A.; THOMAS, C.M.; HAMMOND-KOSACK, K.E.; BALINT-KURTI, P.J.; JONES, J.D.G. Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. **Science**, v. 266, p. 789–793, 1994.
- JONES, N.; OUGHAM, H.; THOMAS, H.. Markers and mapping: We are all geneticists now. **New Phytologist**, v. 137, p.165–177, 1997.
- KAUSHAL, A. et al. **A Rapid Disease Resistance Breeding in Tomato (*Solanum lycopersicum* L.)**. In: GOSAL, S.S.; WANI, S.H. (eds) Accelerated Plant Breeding. v. 2 (Vegetable Crops). Switzerland: Springer Nature Cham, , 2020. p. 17-55.
- KAWCHUK, L.M. et al. Tomato *Ve* disease resistance genes encode cell surface-like receptors. **Proceedings of the National Academy of Sciences USA**, v. 98, p. 6511–6515, 2001.
- KELLY, J.D.; MIKLAS, P.N. The role of RAPD markers in breeding for disease resistance in common bean. **Molecular Breeding**, v. 4, p. 1–11, 1998.
- KIM, C. et al. Application of genotyping by sequencing technology to a variety of crop breeding programs. **Plant Science**, v. 242, p. 14–22, 2016.

- LANDER, E.S. et al. Mapmaker an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. **Genomics**, v. 1, p. 174–181, 1987.
- LAPIDOT, M. et al.. A novel route controlling begomovirus resistance by the messenger RNA surveillance factor pelota. **PLoS Genetics**, v. 11, p. e1005538, 2015.
- LI, J ET AL. Identification and mapping of quantitative resistance to late blight (*Phytophthora infestans*) in *Solanum habrochaites* LA1777. **Euphytica**, v. 179, p. 427–438, 2011.
- LIM, G.T.T. et al. Mapping the *I-3* gene for resistance to *Fusarium* wilt in tomato: application of an *I-3* marker in tomato improvement and progress towards the cloning of *I-3*. **Australasian Plant Pathology**, v. 35, p.671–680, 2006.
- LIM, G.T.T.; WANG, G.-P.; HEMMING, M.N.; MCGRATH, D.J.; JONES, D.A. High resolution genetic and physical mapping of the *I-3* region of tomato chromosome 7 reveals almost continuous microsynteny with grape chromosome 12 but interspersed microsynteny with duplications on *Arabidopsis* chromosomes 1, 2 and 3. **Theoretical and Applied Genetics**, v. 118, p. 57–75, 2008.
- LIN, T et al. Genomic analyses provide insights into the history of tomato breeding. **Nature Genetics**, v. 46, p. 1220–1228, 2014.
- LINNARSSON, S. Recent advances in DNA sequencing methods–general principles of sample preparation. **Experimental Cell Research**, v. 316, p. 1339–1343, 2010.
- LOPES, R. et al. Marcadores moleculares dominantes (RAPD e AFLP). **Biotecnologia Ciência & Desenvolvimento**, v. 29, p. 56–60, 2012.
- MACHADO, M.R. **Eficiência de novas fontes de resistência em tomateiro contra diferentes espécies de Begomovirus bipartidos e localização cromossômica do locus *tcm-1***. Dissertação de Mestrado. Universidade de Brasília (UnB), 2013.
- MANLY, K.F.; CUDMORE, J.R.; R.H.; MEER, J.M. Map Manager QTX, cross-platform software for genetic mapping. **Mammalian Genome**, v. 12, p. 930–932, 2001.
- MARKERT, C.L.; MOLLER, F. Multiple forms of enzymes: Issue, ontogenetic and species-specific patterns. **Proceedings of the National Academy of Sciences USA**, v. 45, p. 753–763, 1959.
- MARTIN, G.B. et al. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. **Science**, v. 262, p. 1432–1436, 1993.
- MEDINA, R., LÓPEZ, S. M., FRANCO, M. E., ROLLAN, C., RONCO, B. L., SAPARRAT, M. C., BALATTI, P. A. A survey on occurrence of *Cladosporium fulvum* identifies race 0 and race 2 in tomato-growing areas of Argentina. **Plant Disease**, v. 99, p. 1732–1737, 2015.
- MEDINA-FILHO, H. P. **An electrophoretic variant as a tool for breeding tomatoes for nematode resistance**. Ph.D. Thesis. University of California, 1980.
- MENDA, N.; STRICKLER, S.R.; MUELLER, L.A. Advances in tomato research in the post-genome era. **Plant Biotechnology**, v. 30, p. 243–256, 2013.
- MICHELMORE, R. W.; PARAN, I.; KESSELI, R.V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. **Proceedings of the National Academy of Sciences USA**, v. 88, p. 9828–9832, 1991.
- MILACH, S.C.K. Marcadores de DNA. **Biotecnologia, Ciência & Desenvolvimento**, v. 5, p. 14–17, 1998.
- MOHAN, M. et al. Genome mapping, molecular markers and marker-assisted selection in crop plants. **Molecular Breeding**, v.3, p. 87–103, 1997.

- MULLIS, K.B.; FALOONA, F.A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. **Methods in Enzymology**, v. 155, p. 335–350, 1987.
- NOVAIS, C.M.; PIRES-ALVES, M.; SILVA, F.F. PCR em tempo real. **Revista Biotecnologia Ciência e Desenvolvimento**, v. 33 p. 10–13, 2004.
- OLIVEIRA, A.S.; BOITEUX, L.S.; KORMELINK, R.; RESENDE, R.O. The *Sw-5* gene cluster: Tomato breeding and research toward orthotospovirus disease control. **Frontiers in Plant Science**, v. 9, p. 1055, 2018.
- ORSI, I.; MALATRASI, M.; BELFANTI, E.; GULLÌ, M.; MARMIROLI, N. Determining resistance to *Pseudomonas syringae* in tomato, a comparison of different molecular markers. **Molecular Breeding**, v. 30, p. 967–974, 2012.
- PANTHEE, D.R.; CHEN F. Genomics of fungal disease resistance in tomato. **Current Genomics**, v. 11, p. 30–39, 2010.
- PARAN, I.; MICHELMORE, R.W. Development of reliable PCR based markers linked to downy mildew resistance genes in lettuce. **Theoretical and Applied Genetics**, v. 85, p. 985–993, 1993.
- PARMAR, P.; SUBRAMANIAN, R.B. PCR based method for testing *Fusarium* wilt resistance of tomato. **African Journal of Basic and Applied Sciences**, v. 3, p. 219–222, 2011.
- PATERSON, A.H. Making genetic maps. In: A.H. Paterson (ed.), **Genome Mapping in Plants**, p. 23–39. R. G. Landes Company, San Diego, California; Academic Press, 1996. Austin, Texas.
- REBRIKOV, D.V.; TROFIMOV, D.Y. Real-time PCR: A review of approaches to data analysis. **Applied Biochemistry and Microbiology**, v. 42, p. 455–463, 2006.
- RICK, C.M.; FOBES, J. Association of an allozyme with nematode resistance. **Report of the Tomato Genetics Cooperative**, v. 24, p. 25, 1974.
- RICK, C.M.; YODER, J.I. Classical and molecular genetics of tomato: Highlights and perspectives. **Annual Review of Genetics**, v. 22, p. 281–300, 1988.
- ROSSI, M. et al. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. **Proceedings of the National Academy of Sciences USA**, v. 95, p. 9750–9754, 1998.
- ROTHAN, C.; DIOUF, D.; CAUSSE, M. Trait discovery and editing in tomato. **The Plant Journal**, v. 97, p. 73–90, 2019.
- SALIBA-COLOMBANI, V.; CAUSSE, M.; GERVAIS, L.; PHILOUZE, J. Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. **Genome**, v. 43, p. 29–40, 2000.
- SALMERON, J.M et al. Tomato *Prf* is member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. **Cell**, v. 86, p. 123–133, 1996.
- SANGER, F.; NICKLEN, S.; COULSON, A.R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences USA**, v. 74, p. 5463–5467, 1977.
- SAYED, H.; KAYYAL, H.; RAMSEY, L.; CECCARELLI, S.; BAUM, M. Segregation distortion in doubled haploid lines of barley (*Hordeum vulgare* L.) detected by simple sequence repeat markers. **Euphytica**, v. 225, p. 265–272, 2002.
- SEMAGN, K.; BABU, R.; HEARNE, S.; OLSEN, M. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. **Molecular Breeding**, v. 33, p. 1–14, 2014.
- SHENDURE, J; JI, H. Next-generation DNA sequencing. **Nature Biotechnology**, v. 26, p. 1135–1145, 2008.

- SHIRASAWA, K; HIRAKAWA, H. DNA marker applications to molecular genetics and genomics in tomato. **Breeding Science**, v. 63, p. 21–30, 2013.
- SIM, S.C. et al. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. **PLoS One**, v. 7, p. e40563, 2012.
- SIMONS, G. ET AL.. Dissection of the *Fusarium I2* gene cluster in tomato reveals six homologs and one active gene. **Plant Cell**, v. 10, p. 1055–1068, 1998.
- SOLLER, M.; BECKMANN, J.S. Genetic polymorphism in varietal identification and genetic improvement. **Theoretical and Applied Genetics**, v. 67, p. 25–33, 1983.
- SPASSOVA, M.I. et al.. The tomato gene *Sw-5* is a member of the coiled coil, nucleotide binding, leucine-rich repeat class of plant resistance genes and confers resistance to TSWV in tobacco. **Molecular Breeding**, v.7, p.151–161, 2001.
- SPOONER, D.M.; PERALTA, I.E.; KNAPP, S. Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes [*Solanum* L. section *Lycopersicon* (Mill.) Wettst.]. **Taxon**, v. 54, p. 43–61, 2005.
- STAM, P. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. **The Plant Journal**, v. 3, p. 739–744, 1993.
- SUN, X.; LIU, D.; ZHANG, X.; LI, W.; LIU, H.; HONG, W.; ZHENG, H. SLAF-seq: An efficient method of large-scale *de novo* SNP discovery and genotyping using high-throughput sequencing. **PLoS One**, v. 8, p. e58700, 2013.
- TANKSLEY, S.D. Mapping polygenes. **Annual Review of Genetics**, v. 27, p. 205–233, 1993
- TANKSLEY, S.D.; FULTON, T.M. Dissecting quantitative trait variation – examples from the tomato. **Euphytica**, v. 154, p. 365–370, 2007
- TANKSLEY, S.D. et al. High density molecular linkage maps of the tomato and potato genomes **Genetics**, v. 132, p. 1141–1160, 1992
- THOMAS, C.M. et al. Characterization of the tomato *Cf-4* gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in *Cf-4* and *Cf-9*. **The Plant Cell**, v. 9, p. 2209–2224, 1997.
- TOMATO GENOME SEQUENCING CONSORTIUM. The tomato genome sequence provides insights into fleshy fruit evolution. **Nature**, v. 485, p. 635–641, 2012.
- TOMATO GENOME SEQUENCING CONSORTIUM. Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. **The Plant Journal**, v. 80, p. 136–148, 2014.
- TOMITA, R. et al. Genetic basis for the hierarchical interaction between *Tobamovirus* spp. and *L* resistance gene alleles from different pepper species. **Molecular Plant Microbe Interactions**, v. 24, p. 108–117, 2011.
- VERLAAN, M.G. et al. Chromosomal rearrangements between tomato and *Solanum chilense* hamper mapping and breeding of the TYLCV resistance gene *Ty-1*. **The Plant Journal**, v. 68, p. 1093–1103, 2011.
- VERLAAN, M.G. et al. The tomato yellow leaf curl virus resistance genes *Ty-1* and *Ty-3* are allelic and code for DFDGD-class RNA-dependent RNA polymerases. **PLoS Genetics**, v. 9, e1003399, 2013.
- VOS, P. et al. AFLP: A new technique for DNA fingerprinting. **Nucleic Acids Research**, v. 23, p. 4407–4414, 1995.
- VOS, P. et al. The tomato *Mi-1* gene confers resistance to both root-knot nematodes and potato aphids. **Nature Biotechnology**, v. 16, p. 1365–1369, 1998.

- WANG, H. et al. Molecular mapping of hypersensitive resistance from tomato 'Hawaii 7981' to *Xanthomonas perforans* race T3. **Phytopathology**, v. 101, p.1217–1223, 2011.
- WANG, Y. et al. Application of whole genome resequencing in mapping of a Tomato yellow leaf curl virus resistance gene. **Scientific Reports**, v. 8, p. 9592, 2018.
- WATSON, J.D.; CRICK, F.H. A structure for deoxyribose nucleic acid. **Nature**, v. 171, p. 737–738, 1953.
- WELSH, J.; MCCLELLAND, M. Fingerprinting genomes using PCR with arbitrary primers. **Nucleic Acids Research**, v. 18, p. 7213–7218, 1990.
- WILLIAMS, J.G.K.; KUBELIK, A.; LIVAK, K.J.; RAFALSKI, J.A.; TINGEY, S.V. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Research**, v.18, p. 6531–6535, 1990.
- WILLIAMSON, V.M.; COLWELL, G. Acid phosphatase-1 from nematode resistant tomato: Isolation and characterization of its gene. **Plant Physiology**, v. 97, p. 139-146, 1991.
- WIN, K.T.; VEGAS, J.; ZHANG, C.; SONG, K.; LEE, S. QTL mapping for downy mildew resistance in cucumber via bulked segregant analysis using next-generation sequencing and conventional methods. **Theoretical and Applied Genetics**, v. 130, p. 199–211, 2017.
- XU, Y.; ZHU, L.; XIAO, J.; HUANG, N.; MCCOUCH, S.R. Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). **Molecular and General Genetics**, v. 253, p. 535–545, 1997.
- YAMAGUCHI, H. et al. An NB-LRR gene, TYNBS1, is responsible for resistance mediated by the *Ty-2* begomovirus resistance locus of tomato. **Theoretical and Applied Genetics**, v.131, p. 1345–1362, 2018.
- YANG, W.; FRANCIS, D.M. Marker-assisted selection for combining resistance to bacterial spot and bacterial speck in tomato. **Journal of the American Society for Horticultural Sciences**, v. 130, p. 716–721, 2005.
- YANG, X. et al. Fine mapping of the tomato yellow leaf curl virus resistance gene *Ty-2* on chromosome 11 of tomato. **Molecular Breeding**, v. 34, p. 749–760, 2014.
- YOUNG, N.D. QTL mapping and quantitative disease resistance in plants. **Annual Review of Phytopathology**, v. 34, p. 479–501, 1996
- YOUNG, N.D.; TANKSLEY, S.D. Restriction fragment length polymorphisms maps and the concept of graphical genotypes. **Theoretical and Applied Genetics**, v. 77, p. 95–101, 1989.
- ZAMIR, D. Improving plant breeding with exotic genetic libraries. **Nature Reviews Genetics**, v. 2, p. 983–989, 2001.
- WANG, Z; GERSTEIN, M; SNYDER, M. RNA-Seq: A revolutionary tool for transcriptomics. **Nature Reviews Genetics**, v. 10, p. 57, 2009.
- ZHANG, C. et al. The *Ph-3* gene from *Solanum pimpinellifolium* encodes CC-NBS-LRR protein conferring resistance to *Phytophthora infestans*. **Theoretical and Applied Genetics**, v. 127, p. 1353–1364, 2014
- ZHAO, T et al. Mapping and candidate gene screening of tomato *Cladosporium fulvum*-resistant gene *Cf-19*, based on high-throughput sequencing technology. **BMC Plant Biology**, v.16, p. 51, 2016.